Norcoclaurine synthase: structural studies, enzyme engineering and the biocatalytic synthesis of novel alkaloids

Birkbeck, University of London

Norcoclaurine Synthase: Structural Studies, Enzyme Engineering, and the Biocatalytic Synthesis of Novel Alkaloids

Rebecca Roddan

2021

Thesis submitted to Birkbeck, University of London for the degree of Doctor of Philosophy in Chemical Biology

Declaration

The cloning of all Δ29*Tf*NCSs (wild-type and variants) and wild-type Δ33*Tf*NCS was carried out by Dr. Benjamin Lichman. The expression of single point variants of Δ29*Tf*NCS was performed by Dr. Daniel Méndez-Sánchez. The cloning of Δ33*Tf*NCS-M97V and the *Nn*NCSs was carried out by Professor John Ward. The cloning, expression and buffer exchange of Δ33*Tf*NCS-L72A, L72V, F80W and F80Y was performed by Yu Wang. Methyltransferase enzymes and co-factor generation and degradation enzymes were provided as glycerol stocks or lyophilised enzyme by Dr. Fabiana Subrizi, Dr. Michael Richter and Prof. Jennifer Andexer. Helena Philpott devised the GC method used for the chiral separation of 2-methylbutanal. Crystallographic investigations were carried out in collaboration with Dr. Altin Sula.

A version of Section 1.7 has been published: R. Roddan, J. M. Ward, N. H. Keep and H. C. Hailes, *Curr. Opin. Chem. Biol.*, 2020, **55**, 69–76.

A version of Chapter 3 has been published: R. Roddan, G. Gygli, A. Sula, D. Méndez-Sánchez, J. Pleiss, J. M. Ward, N. H. Keep and H. C. Hailes, *ACS Catal.*, 2019, **9**, 9640–9649.

A version of Chapter 5 has been published: R. Roddan, A. Sula, D. Méndez-Sánchez, F. Subrizi, B. R. Lichman, J. Broomfield, M. Richter, J. N. Andexer, J. M. Ward, N. H. Keep and H. C. Hailes, *Commun. Chem.*, 2020, **3**, 1–10.

I, Rebecca Roddan, confirm that this thesis presents work that is my own and that where information has been derived from other sources, this has been indicated.

Abstract

Norcoclaurine synthase (NCS) is a biocatalyst, involved in plant alkaloid biosynthesis. In plants, NCS catalyses a stereoselective Pictet-Spengler reaction between dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA), to give (*S*)-norcoclaurine, the first committed intermediate in benzylisoquinoline alkaloid (BIA) biosynthesis. Previous work has shown that a range of aldehydes and ketones are accepted by NCS, in place of 4-HPAA, to generate single isomer tetrahydroisoquinolines (THIQs). The THIQ moiety is found in many biologically active molecules and NCS provides a facile route towards novel analogues.

The aim of this project was to rationalise and extend the substrate scope of NCS, in particular towards α-substituted aldehydes. Despite previous reports, a range of α-methyl substituted aldehydes were shown to be accepted by NCS. A kinetic resolution of the aldehyde was also observed, with the (*R*)-enantiomer preferentially accepted, leading to THIQ products with two well-defined stereocentres. Single point variants of NCS were shown to improve activities compared with the wild-type enzyme. One variant, M97V was particularly promising and lead to the acceptance of the benzaldehydes as substrates and so the single step syntheses of various (1*S*)-aryl-THIQs was achieved. To expand the amine substrate scope of NCS, *N*-methyl-phenylamines were also explored as substrates.

Chemoenzymatic cascades were developed to increase the molecular complexity of the NCS-generated THIQs and crystallographic investigations, involving the co-crystallisation of reaction intermediate mimics helped to rationalise NCS activities. Routes towards (*R*)-THIQs were also explored, involving screening novel variants, NCSs from *Nelumbo nucifera* and another Pictet-Spenglerase, salsolinol synthase.

Acknowledgements

I am incredibly grateful to my supervisors, Professors Nicholas Keep, Helen Hailes and John Ward, for all their expertise and support over the past four years. I really couldn't have imagined a better team to guide me nor a project I would have enjoyed more.

I owe much of the success of the protein crystallography work to Dr. Altin Sula. I have felt lucky to be taught by someone so knowledgeable and I am grateful for all the time you've spent helping me with this project. Thank you to my supervisor, Professor Nicholas Keep, for spending so much time teaching me how to do the data processing.

Thank you also to Dr. Jenny Pritchard, Dr. Alice Dunbabin, Dr. Jianxiong Zhao, Dr. Daniel Méndez-Sánchez, Dr. Yu Wang, and Dr. Fabiana Subrizi for teaching me everything I needed to know in the lab. I am so grateful to them and other past and present members of the Hailes and Tabor groups for their advice, for the many muffin breaks, and for being a wonderful group of people to spend time with.

Thank you to my thesis chair, Dr. Tracey Barrett for your support and guidance on my thesis and to Dr. Claire Bagneris for all the help you've given me in the lab. Thank you also to Dr. Ambrose Cole and Dr. Nikos Pinotsis for their help with x-ray data collection. I am also thankful to Dr. Kersti Karu and Dr. Abil Aliev for their help with mass spectroscopy and NMR spectroscopy.

I am also indebted to Birkbeck College for providing funding and the London Interdisciplinary Biosciences PhD Programme for giving me the opportunity to do a PhD in an incredible research environment.

Thank you to my family and friends for your support throughout and for always believing in me. Last, but not least, Hamish. I am so grateful for your unwavering positivity, motivation and understanding, I truly couldn't have done it without you.

Table of Contents

List of Abbreviations

13-MeTHPB	13-Methyltetrahydroprotoberberine
13-MePB	13-Methylprotoberberine
3,4-DHPAA	3,4-Dihydroxyphenylacetaldehyde
3,4-DHPP	3,4-Dihydroxyphenylpyruvate
4-HPAA	4-Hydroxyphenylacetaldehyde
4-MOPAA	4-Methoxyphenylacetaldehyde
ATP	Adenosine triphosphate
BBE	Berberine Bridge Enzyme
BIA	Benzylisoquinoline alkaloid
C.	*Corydalis*
Cj	*Coptis japonica*
COMT	Catechol *O*-methyltransferase
Cr	*Catharanthus roseus*
d.r.	Diastereomeric ratio
DIS	Deacetylipecoside synthase
DMF	*N,N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
E. coli	*Escherichia coli*
Ec	*Escherichia coli*
e.e.	Enantiomeric excess
e.q.	Equivalents
GC	Gas chromatography
h	Hours
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His-tag	Hexahistidine-tag
HRMS ES+	Positive electrospray high resolution mass spectrometry
Hz	Hertz
IPTG	Isopropyl β-d-1-thiogalactopyranoside
IR	Infrared
IRED	Imine reductase
J	Coupling constant
K_d	Dissociation constant
LB	Lysogeny broth
LCMS	Liquid chromatography mass spectrometry
MAO	Monoamine oxidase
MAT	Methionine adenosyl transferase

MD	Molecular dynamics
MeCN	Acetonitrile
MES	2-(*N*-Morpholino)ethanesulfonic acid
MS	Mass spectrometry
MTAN	Methylthioadenosine nucleosidase
Mx	*Myxococcus xanthus*
NBS	Norbelladine Synthase
NCS	Norcoclaurine Synthase
NMR	Nuclear Magnetic Resonance
Nn	*Nelumbo nucifera*
NOEsy	Nuclear Overhauser Effect Spectroscopy
PB	Protoberberine
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PR	Pathogenesis-related
Ps	*Papaver somiferum*
Rn	*Rattus norvegicus*
RP	Reversed phase
rpm	Revolutions per minute
Rs	*Rauvolfia serpentina*
S. cerevisiae	*Saccharomyces cerevisiae*
SAM	(*S*)-Adenosyl methionine
SDS-PAGE	Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis
SOC media	Super Optimal Broth with Catabolite repression
SS	Salsolinol synthase
STR	Strictosidine synthase
TEA	Triethylamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
THIQ	Tetrahydroisoquinoline
THPB	Tetrahydroprotoberberine
Tf	*Thalictrum flavum*
TLC	Thin-layer chromatography
t_R	Retention time
TRIS	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
WT	Wild type

List of Schemes

List of Figures

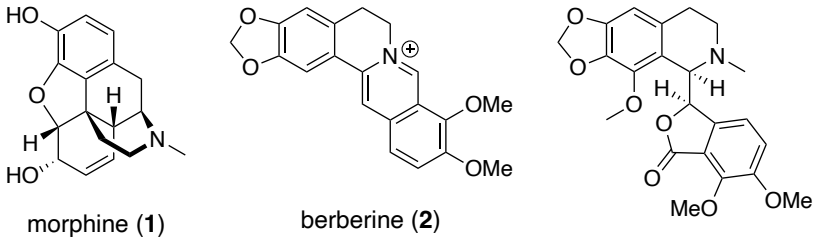
List of Tables

Chapter 1: Introduction

1.1 Background

1.1.1 Benzylisoquinoline alkaloids

Benzylisoquinoline alkaloids (BIAs) are a diverse group of nitrogen-containing plant metabolites, commonly found in the *Ranunculales* order of plants which includes plants such as poppies, meadow grass and buttercups. Over 2,500 different natural BIAs have been identified and are characterised by the common benzylisoquinoline moiety. Despite the *in planta* roles of these metabolites being largely unknown, many BIAs possess potent medicinal properties which have been exploited for many centuries.[1] For example, the opium poppy has been used for its analgesic properties since the Neolithic age, with the widely-known BIA, morphine (**1**) being responsible for these activities. Other examples of widely used BIAs include berberine (**2**), with anti-microbial properties and noscapine (**3**), an antitussive.[2] Many others have been shown to exhibit anti-HIV, anti-tuberculosis and anti-cancer properties.[3,4] Figure 1.1 exemplifies the large structural differences of three medicinally important BIAs.

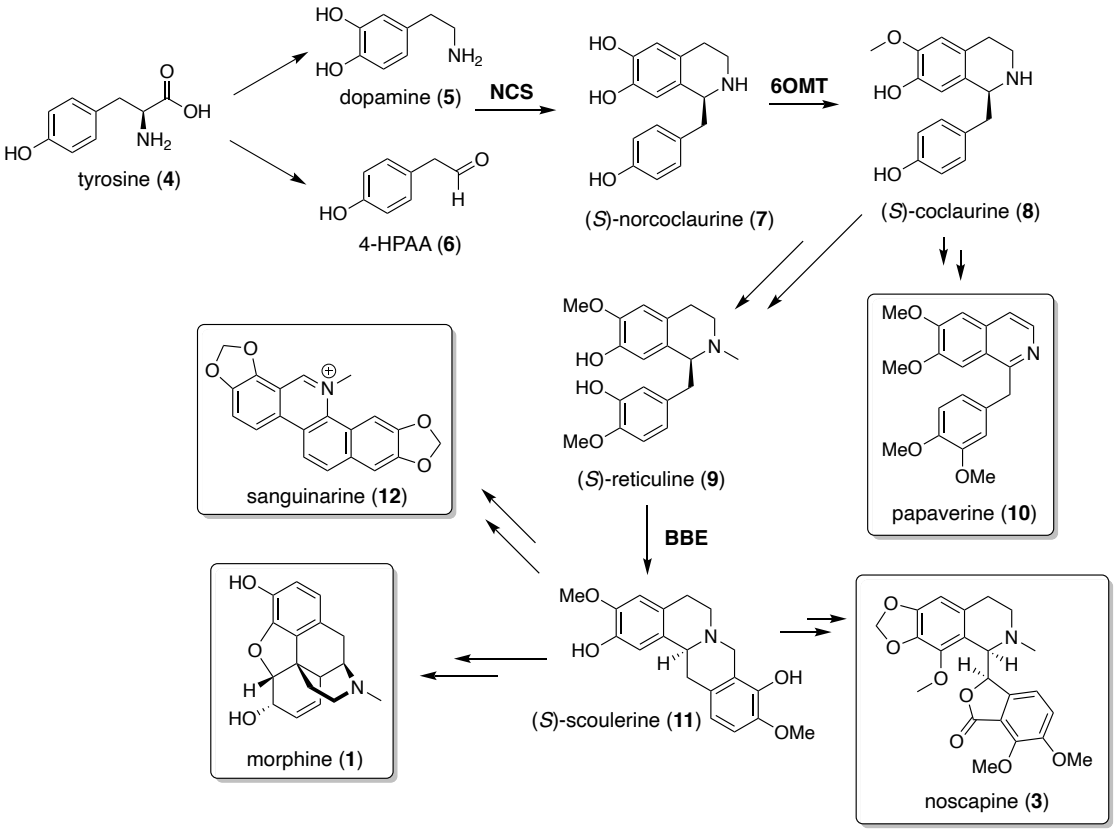


morphine (**1**) berberine (**2**) noscapine (**3**)

Figure 1.1: Examples of biologically active BIAs.

The medicinal properties of BIA-producing plants have been exploited for millennia, leading to recent scientific efforts into the identification of the active components. Currently, around 50% of marketed drugs are based upon natural products,[5] however, isolation of a single metabolite from plants is challenging as it must be separated from other structurally similar molecules and low quantities are produced. Therefore, significant research efforts have been focused on understanding the biosynthetic

pathways involved so that the transformations can be replicated in the laboratory, either by using analogous chemical transformations using the relevant recombinantly-expressed enzymes *in vitro* or biosynthesised via whole-cell production. Due to the pharmaceutical relevance of the BIAs, the full biosynthetic pathways to the most widely known BIAs have been elucidated over a number of years. Indeed, despite the obvious medicinal value of morphine, the biosynthetic pathway was not fully realised until 2015 with the discovery of the reticuline epimerase enzyme.[6,7] The biosynthetic pathways to several BIAs are given in Scheme 1.1.

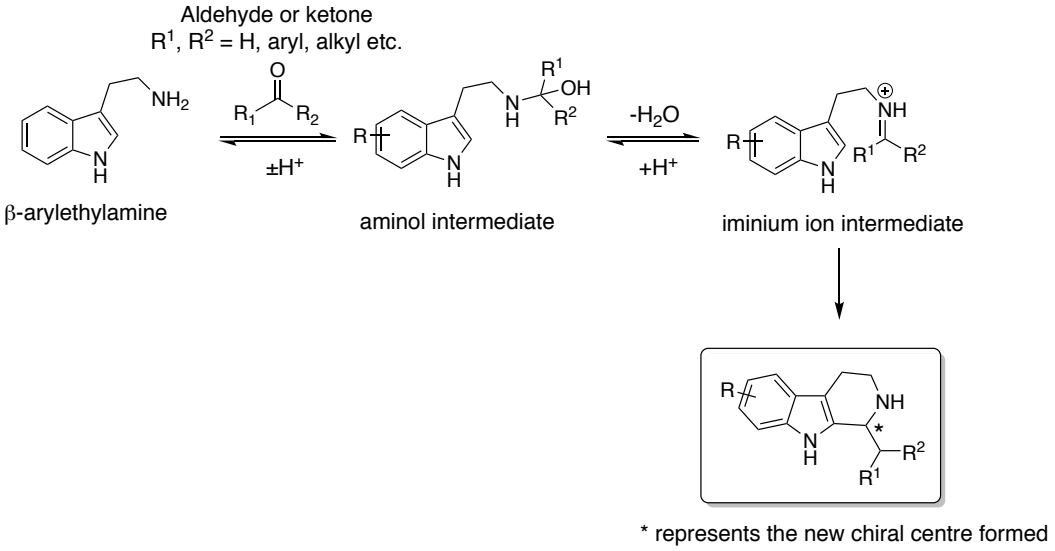


Scheme 1.1: The biosynthetic pathways towards a variety of BIAs. Given on the arrows (in bold) are some of the enzymes responsible (NCS: norcoclaurine synthase, 6OMT: norcoclaurine-6-*O*-methyltransferase, BBE: berberine bridge enzyme). A double arrow signifies where multiple enzymatic reactions are responsible for conversion. Scheme adapted from Hagel *et al.*[1]

1.1.2 The Pictet-Spengler reaction

The first-committed intermediate in the biosynthetic pathway to BIAs is currently known to be (*S*)-norcoclaurine (**7**), formed via a stereoselective Pictet-Spengler reaction

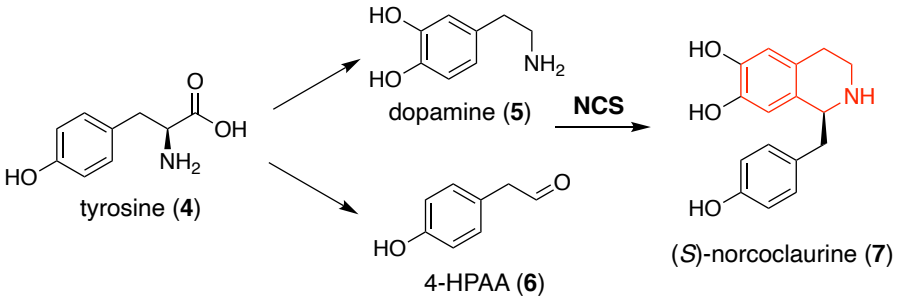
between two tyrosine (**4**) derived molecules, dopamine (**5**) and 4-hydroxyphenylacetaldehyde (**6**) (Scheme 1.1). The Pictet-Spengler reaction was first reported by Amè Pictet and Theodor Spengler in 1911 and can be considered as an intramolecular version of the Mannich reaction.[8] The reaction involves an intramolecular cyclisation reaction between a β-phenylethylamine and an aldehyde or ketone. First, the amine and the aldehyde condense to give an iminium ion, followed by electrophilic addition of the aromatic system to the iminium ion, with aromaticity regained by deprotonation. The reaction is advantageous for the generation of significant molecular complexity in a single reaction step as a new ring and two bonds (C-C and C-N) are formed. The reaction is generally performed chemically under acidic conditions and a representative scheme of the reaction is given below (Scheme 1.2).



Scheme 1.2: The general reaction pathway of the Pictet-Spengler reaction.

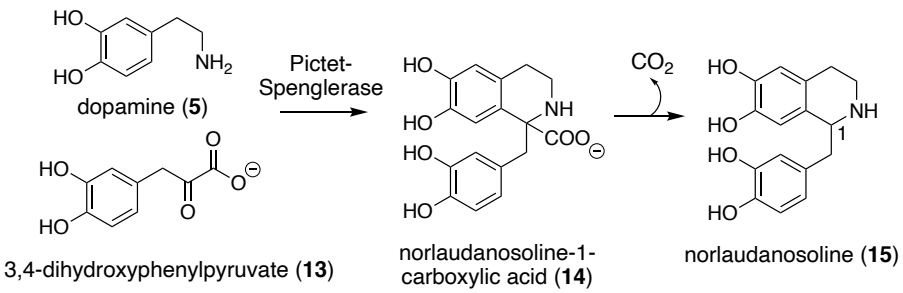
1.2 Norcoclaurine Synthase

All BIAs have the common precursor, (*S*)-norcoclaurine (**7**), which is generated by a Pictet-Spengler reaction between two tyrosine-derivatives, dopamine (**5**) and 4-HPAA (**6**) (Scheme 1.3). The selectivity is generated by the Pictet-Spenglerase, norcoclaurine synthase (NCS). The role of a Pictet-Spengler reaction in BIA biosynthesis was initially hypothesised by Winterstein *et al.*[9] due to the presence of the tetrahydroisoquinoline (THIQ) moiety.



Scheme 1.3: The biosynthesis of (*S*)-norcoclaurine. The THIQ moiety generated is shown in red.

Initial feeding studies for the discovery of the first committed intermediate in BIA biosynthesis, involved the BIA-producing opium poppy (*Papaver orientale*). Isotopically-labelled ^{14}C-tyrosine (**4**) was incorporated into both aromatic rings and labelled ^{14}C,^{3}H-dopamine (**5**) was found to be incorporated into norlaudanosoline (**15**) and norlaudanosoline-1-carboxylic acid (**14**).[10] Therefore, it was proposed that a Pictet-Spengler reaction between dopamine (**5**) and 3,4-hydrohydroxyphenylpyruvate (**13**) occurred, followed by a decarboxylation of norlaudanosoline-1-carboxylic acid (**14**) to give norlaudanosoline (**15**) (Scheme 1.4). Complementary incorporation experiments into *P. somniferum* have also been performed by Battersby *et al.*[11]



Scheme 1.4: Proposed biosynthetic pathway to give norlaudanosoline. This route was hypothesised by Wilson *et al.*[10] based upon feeding studies.

The Pictet-Spenglerase involved in this first step was identified in 1981 by Rueffer *et al.*[12] Kinetic isotope effect experiments on the enzyme-containing extracts of ten different BIA-producing plants identified an enzyme that formed norlaudanosoline (**15**) by a Pictet-Spengler reaction between dopamine (**5**) and 3,4-dihydroxyphenylacetaldehyde (3,4-DHPP); i.e. the aldehyde rather than the α-ketoacid

(**13**) which was previously proposed by Wilson *et al*.[10] No incorporation of 3,4-hydrohydroxyphenylpyruvate (**13**) was observed, leading to the assumption that the condensation observed may be chemical rather than enzymatic. Feeding studies showed the acceptance of 4-HPAA by the enzyme and the stereochemistry of the C-1 centre (Scheme 1.4) was realised by circular dichroism of isolated norlaudanosoline (**15**) compared with standards of known absolute configuration.[12]

Further feeding studies by Schumacher *et al.* in 1983 confirmed the lack of incorporation of phenylpyruvates and the incorporation of 4-HPAA (**6**) and 3,4-DHPP (**13**) as the carbonyl substrate.[13] However, analysis of the results by Stadler *et al*. suggested that the two tyrosine-derivatives incorporated differ from one another as dopamine was not incorporated into the benzylic portion of norlaudanosoline (**15**) ; i.e. the biosynthesis of one of the substrates is not via the same biosynthetic pathway as the other.[14] Feeding studies involving labelled (*R*) or (*S*)-norcoclaurine showed the specific incorporation of the (*S*)-enantiomer into cocolaurine (**8**), reticuline (**9**) and thebaine, an opiate alkaloid which is structurally similar to morphine (**1**). This therefore suggested that (*S*)-norcoclaurine (**7**) was the first committed precursor in BIA biosynthesis and the Pictet-Spenglerase involved was renamed (*S*)-norcoclaurine synthase (NCS).[14]

1.2.1 Isolation and characterisation of norcoclaurine synthases

The first wild-type NCS was isolated from a member of the poppy family, *Eschscholzia tenuifolia* in 1983.[13] The enzyme was found to be a 15.5 kDa protein with pH and temperature optima of 7.8 and 40 ^{o}C respectively. A non-enzymatic, racemic, phosphate-catalysed background Pictet-Spengler reaction was also noted, the full potential of which was not realised until recently.[15] In 2001, Samanani *et al.* isolated NCSs from three different BIA-producing plants; *Thalictrum flavum* (common meadow-rue), *Eschscholzia californica* (California poppy) *and Papaver somniferum* (opium

poppy). It was found that the optimal temperature for NCS activity for the three enzymes was between 42 and 55 ^{o}C and the optimal pH for the reactions was between 6.5 and 7. The enzyme also exhibited sigmoidal saturation kinetics for dopamine with 4-HPAA at saturating substrate concentrations, indicating that there is positive co-cooperativity between substrate binding sites, i.e., the binding of 4-HPAA increases dopamine binding. Based upon analysis of other enzymes with similar substrate binding, the authors suggested that NCS has a regulatory role in BIA biosynthesis and is involved in controlling the flux of the biosynthetic pathway.[16]

The genes encoding for *Tf*NCS (isolated from *Thalictrum flavum*) were identified in 2004, leading to the first recombinant expression of the enzyme in *E. coli*. Constructs with *N*-terminal truncations (10 or 19) and a *C*-terminal hexahistidine tag (His-tag) were successfully expressed and shown to retain typical NCS activity. The first 19 *N*-terminal residues were identified to be a signal peptide, suggesting that NCS is located in a subcellular compartment.[17] Indeed, studies of the subcellular localisation of dopamine and related alkaloids in *Papaver bracteatum* have indicated that dopamine is predominantly located within a 'vacuolar' compartment.[18] The sequence similarity of NCS to other proteins was also explored. Significant similarity to the Bet v 1 allergen and PR10 proteins was noted, however, there was no similarity to any other known Pictet-Spenglerases such as strictosidine synthase and deacetylipecoside synthase.[17,19]

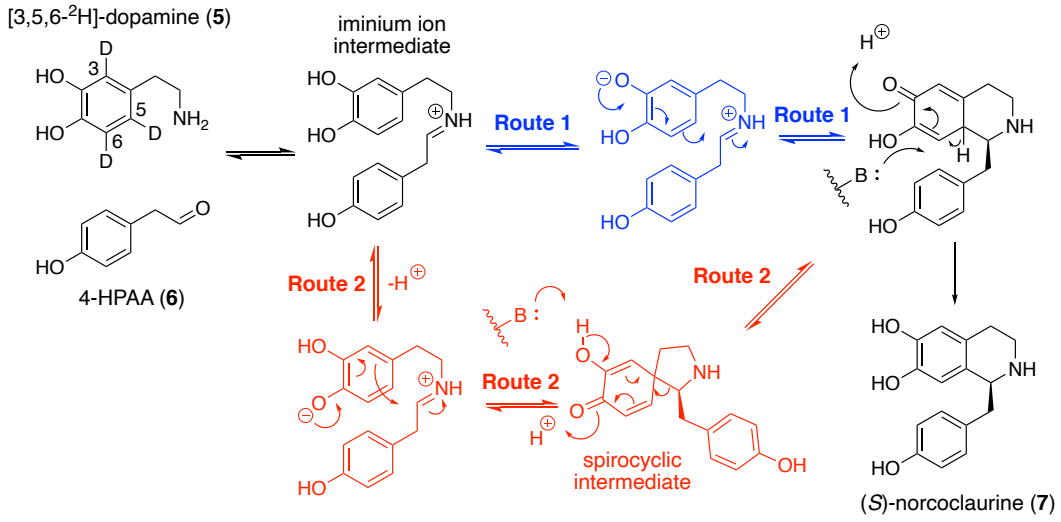
To date, NCSs have been identified in several BIA-producing plants, expressed recombinantly and the activities studied.[20] The two variants that are mostly widely explored are *Tf*NCS and *Cj*NCS2 (isolated from *Coptis japonica*).[21,22] A multiple sequence alignment of *Cj*NCS2 and *Tf*NCS is shown in Figure 9.56.

1.3 Structural and mechanistic investigations of norcoclaurine synthase

1.3.1 Preliminary kinetics experiments

Initial substrate saturation kinetics experiments of *Tf*NCS performed by Samanani *et al*. indicated that hyperbolic saturation kinetics are observed upon 4-HPAA (**6**) addition, i.e., saturation of the enzyme is observed at high substrate concentrations, in the absence of the other substrate. Sigmoidal saturation kinetics were observed with dopamine, indicating positive cooperativity, i.e., the affinity of binding of dopamine (**5**) to another binding site is increased upon binding of 4-HPAA (**6**). These results therefore, suggested that **6** binds to the NCS active site before dopamine (**5**) and that there are possibly multiple ligand binding sites.[16,23]

The first complete purification and mechanistic investigations of *Tf*NCS were performed by Luk *et al*. in 2007. Recombinant protein was expressed in *E. coli* and purified by affinity chromatography.[24] Circular dichroism (CD) and ^{1}H-NMR based kinetic assays were used to elucidate the mechanism of action. It was hypothesised that the reaction would proceed via either a direct cyclisation (Route 1, Scheme 1.5) or a spirocyclic intermediate (Route 2, Scheme 1.5).



Scheme 1.5: Two possible NCS mechanisms proposed by Luk *et al*. (redrawn from Figure 3).[24] Route 1 (blue) proceeds via a quinone intermediate which is deprotonated to regain aromatisation. Route 2 (red) proceeds via a spirocyclic intermediate which undergoes a semi-pinacol rearrangement to give the quinone intermediate, followed by deprotonation (black).

Various analogues of dopamine were tested as substrates and the *para*-hydroxyl was shown to be non-essential for a successful enzymatic reaction, leading to the conclusion that the reaction proceeds via Route 1 (Scheme 1.5). In both proposed routes, deprotonation of 5-H in dopamine (**5**) is essential to regain aromaticity and generate the product. A kinetic isotope effect was observed when comparing reaction rates between using unlabelled dopamine (**5**) and [3,5,6-H^2]-**5**, thus suggesting that the rearomatisation step is partially rate determining.[24] Subsequent product inhibition kinetics experiments also suggested that the conformations of the enzyme upon aldehyde or product binding differ, eluding to structural changes of the enzyme during the mechanistic process. The presence of a CD signal when monitoring the progress of the reaction also confirmed that the NCS reaction is stereoselective, as previously observed by Stadler *et al*.[14]

1.3.2 Nuclear magnetic resonance studies

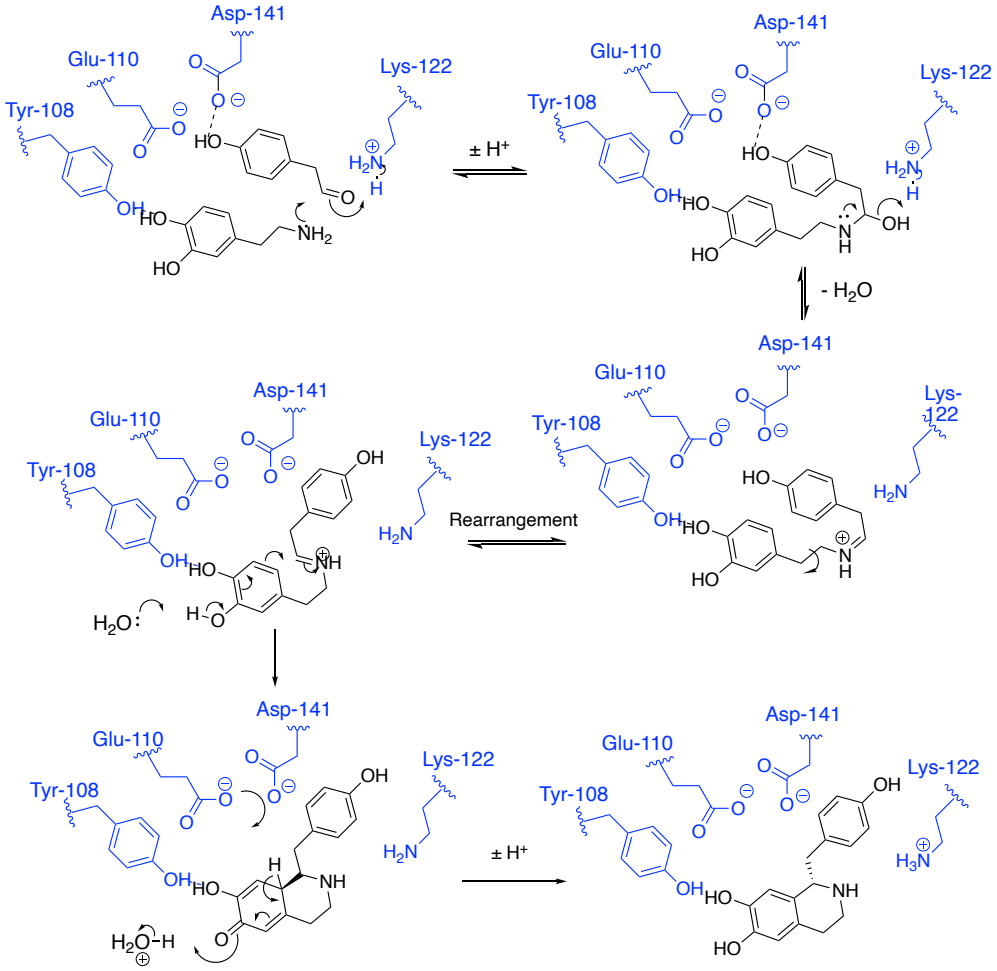
The first structural studies of *Tf*NCS were performed by nuclear magnetic resonance (NMR) spectroscopy. NCS had previously been shown to have high sequence similarity with the Bet v 1 allergen family of proteins.[25] Comparison of the CD spectra of *Tf*NCS and a Bet v 1 protein (PDB: 1BV1) suggested that the two proteins contained similar proportions of the same secondary structure elements. Therefore, a Bet v 1 protein was used as the template in structure determination.[26] The enzyme was recombinantly expressed with a 19 residue *N*-terminal truncation (Δ19*Tf*NCS), however, SDS-PAGE analysis of the isolated protein showed the presence of the desired construct with an additional degradation product, with a 29 residue *N*-terminal truncation (Δ29*Tf*NCS).[25] Enzymatic activity was unaffected by this additional truncation, so NMR studies were performed with Δ29*Tf*NCS. The results confirmed that NCS is structurally similar to PR10 and Bet v 1 proteins. Substrate binding assays were performed with dopamine (**5**) and methyl 4-hydroxyphenylacetate or 2-(4-hydroxyphenyl)ethanol (i.e. methyl ester

and alcohol analogues of the natural substrate, 4-HPAA (**6**)). **6** was not used due to issues of oxidative sensitivity, meaning that the substrate could not be obtained in sufficient purity.

Upon dopamine binding, conformational shifts of residues 69-72 and 150-155 were observed and binding of both aldehyde analogues resulted in significant conformation changes across the entire protein. Unfortunately, this meant that the key active site residues involved in NCS catalysis, and the locations of substrate binding sites could not be identified. The K_d value of dopamine was found to be high (in the mM range), meaning that there is low affinity to the enzyme. Dopamine in BIA-producing plants has been shown to accumulate in vacuolar compartments, which could compensate for the low substrate binding affinity observed. Oligomerisation of the enzyme was also noted, with dimerisation observed at concentrations over 10 μM. Enzymatic activity was retained in the monomeric state during NMR studies at < 10 μM.[26]

1.3.3 Initial crystallographic studies

Although NMR studies had successfully identified the overall fold of NCS, the apparent dynamic nature of the mechanistic process hindered the identification of the key active site residues involved. The first crystallographic investigations of *Tf*NCS were performed by Ilari *et al*.[27] with Δ29*Tf*NCS. A structure was gained at 2.1 Å (PDB: 2VQ5) with the enzyme complexed with dopamine (**5**), and a non-productive 4-HPAA (**6**) analogue, 4-hydroxybenzaldehyde, thus leading to the identification of catalytically relevant active site residues. The structure confirmed the PR10 fold of the enzyme and identified a 23.4 Å catalytic tunnel. The two ligands were built into the positive difference density observed in the active site, suggesting that the aldehyde binds first followed by dopamine (**5**). The 'aldehyde-first' mechanism proposed is given in Scheme 1.6.



Scheme 1.6: The 'aldehyde-first' mechanism. The mechanism was proposed based upon a crystal structure of *Tf*NCS with dopamine (**5**) and a non-productive 4-HPAA (**6**) analogue bound in the enzyme active site. Redrawn from Figure 6 of Ilari *et al*.[27]

In the proposed mechanism, the aldehyde is held in the enzyme active site by a hydrogen-bonding interaction between the aldehyde carbonyl and K122. Dopamine (**5**) then binds to the active site and there is a nucleophilic attack of the amino group onto the aldehyde carbonyl, leading to condensation of the two substrates and iminium ion formation. The iminium ion then reorientates so that it is in the correct orientation for intramolecular cyclisation via electrophilic attack of the catechol aryl ring to the iminium carbon. The 5-H proton is then abstracted by E110, leading to product formation.

1.3.4 Confirming the NCS mechanism

Further analysis of the co-crystallised structure of *Tf*NCS (PDB: 2VQ5)[27] and computational docking experiments revealed mechanistic flaws with the 'aldehyde-first' mechanism.[28] Fundamentally, the ligands modelled into the active site in the paper published by Ilari *et al*. did not fit the electron density map convincingly. Some enzymes are capable of retaining and / or exhibiting altered catalytic activity when crystalline, so with the first NCS structure (PDB: 2VQ5), it is possible that a combination of poor ligand occupancy, multiple ligand orientations or a mixture of different reaction intermediates was observed, thus resulting in poor difference density in the active site.[29] The R-factors after refinement were also poor. The 'aldehyde-first' mechanism assumes that the aldehyde is buried in the active site which would not be possible for bulky aldehydes such as citronellal and napthylacetaldehyde for which NCS activity is observed.[21,22] Many of the aldehydes capable of turnover by NCS would furthermore not be able to form the interaction with D141 that was suggested as they lack the phenolic hydroxyl group. In addition, there was no evidence that the suggested rearrangement of the iminium intermediate would be possible, water is not basic enough to deprotonate the C-3 phenol of dopamine (and so a basic active site residue must be needed) and the hydrogen bond between the quinone intermediate and Y108 would not provide sufficient stabilization. Indeed, the enzyme is active without the presence of the C-4 phenol so this interaction must not be involved.

Subsequent crystallographic studies were used to support the 'dopamine-first' mechanism proposed from computational docking experiments.[28,30] Initial crystallisation attempts with the Δ29*Tf*NCS construct resulted in poor resolution datasets (3.5 – 4 Å) and it was hypothesised that the long *C*-terminal extension was preventing tight crystal packing.[31] A novel construct was therefore designed, omitting residues 1-33 and 196-210, to give just the globular fold. Enzymatic activity was retained, and this construct was used for crystallographic studies. Instead of attempting to bind substrates, a non-

productive analogue of the iminium ion intermediate of a successful NCS reaction was synthesised and successfully co-crystallised with ΔN33C196*Tf*NCS (PDB: 5NON, 1.8 Å resolution). The mimic used was based upon the reaction between dopamine (**5**) and a non-productive analogue of 4-HPAA (**6**), 4-methoxyphenylacetaldehyde (4-MOPAA) to avoid issues of oxidative-sensitivity. Visualisation of the structure gained and the rationale for the design of the mimic is given in Figure 1.2. Upon mimic co-crystallisation, the tertiary structure of NCS was conserved (PDB: 5NON, beige, Figure 1.2c) compared with the apo structure (PDB: 5N8Q, blue, Figure 1.2c). There were some changes in the positions of amino acid side chains of key active site residues (e.g. E110, M183, D141, Y108, M97).

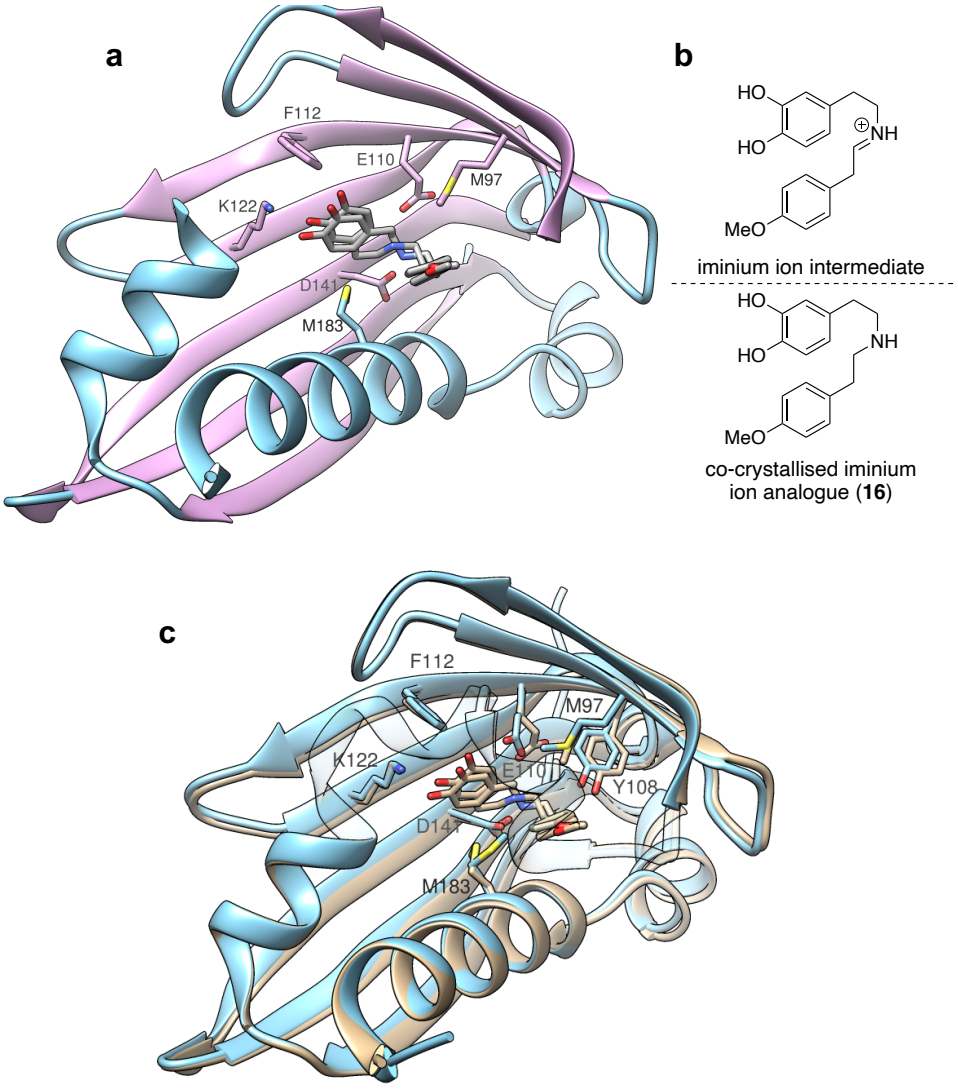
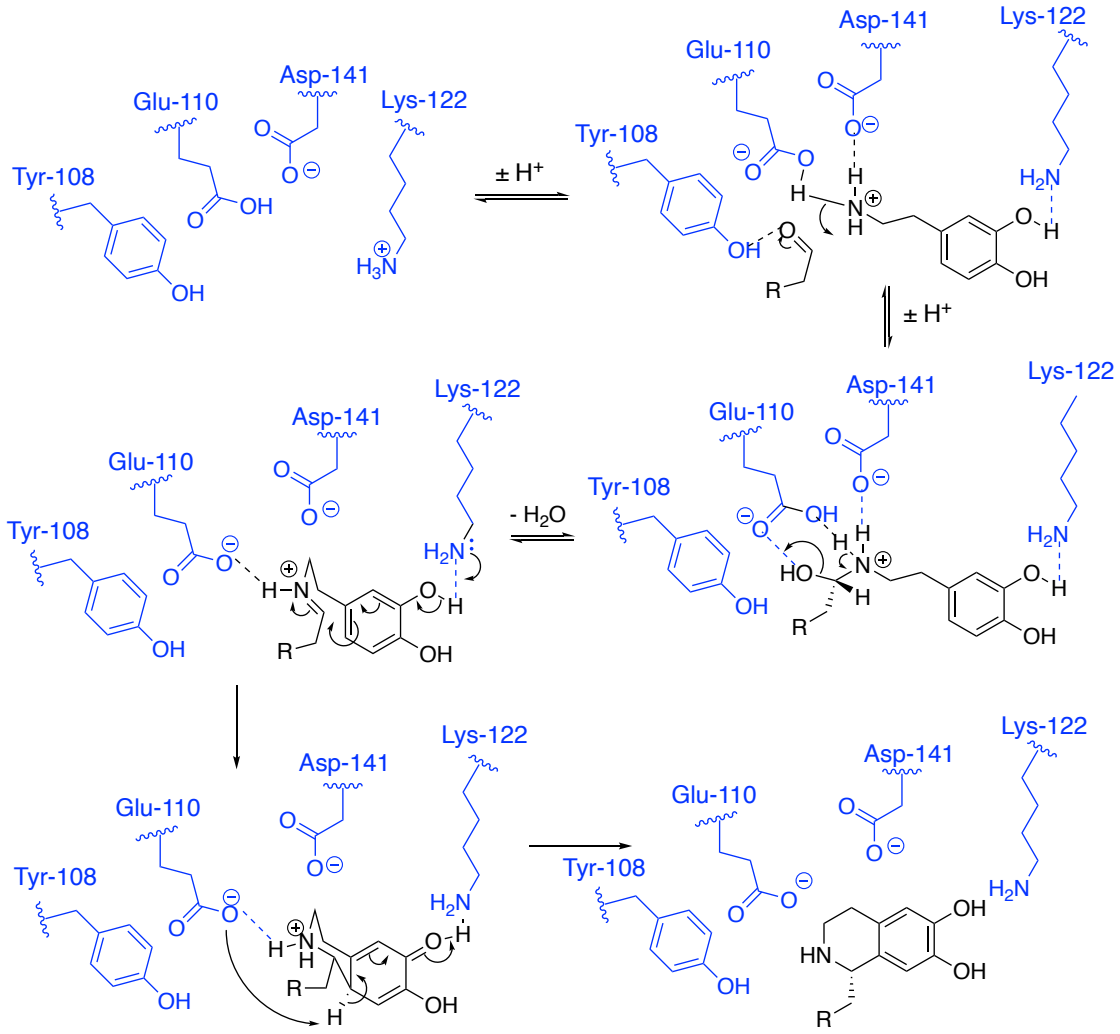


Figure 1.2: Co-crystallised structure of *Tf*NCS with an active site bound reaction intermediate analogue. a) Subunit A of PDB: 5NON was used and the bound mimic is given in

grey.[30] Amino acids 65-79 were removed to aid visualisation of the active site. Image was prepared using USCF Chimera.[32] b) Rationale for mimic design, based upon the iminium ion intermediate of the Pictet-Spengler reaction between dopamine and a non-productive analogue of 4-HPAA (**6**). c) Overlayed apo (PDB: 5N8Q, blue) and holo (PDB: 5NON) structures of Δ33*Tf*NCS.[30]

The mimic-bound structure revealed two different binding modes of the ligand: productive and inhibitory. The productive mode has the *meta*-hydroxyl of the catechol of **16** hydrogen-bonded to K122, whereas in the inhibitory mode, the interaction is observed with the *para*-hydroxyl instead. The *para*-hydroxyl is non-essential for a productive reaction and so this binding mode will not lead to turnover by the enzyme. A revised, 'dopamine-first' mechanism considering these factors was proposed and confirmed through site-specific mutations at key active site residues; K122, E110, D141 and Y108. Further studies of the mechanism and the enantioselectivity of the reaction have been performed using quantum chemical modelling.[33] The investigations involved the two natural substrates, dopamine (**5**) and 4-HPAA (**6**). Both reaction pathways, 'aldehyde-first' or 'dopamine-first', were shown to be accommodated in the enzyme active site but lower energy barriers were associated with the 'dopamine-first' pathway. The selectivity of the reaction was shown to be determined in the final step, whereby the quinone intermediate is deprotonated to yield the THIQ product (Scheme 1.5). In the pathway to give the (*R*)-norcoclaurine, the 4-hydroxybenzyl moiety would be held in an axial position, leading to an unfavourable steric clash with L72. The pathway to (*S*)-norcoclaurine avoids this interaction and, therefore, is favoured. The study also confirmed the key mechanistic roles of K122, D141 and E110. The current proposed mechanism of action of NCS is given in Scheme 1.7.



Scheme 1.7: The currently accepted, 'dopamine-first' mechanism of NCS catalysis. The mechanism was proposed based upon computational docking[28] and crystallographic[30] studies. Hydrogen bonding interactions are shown by dotted lines.

1.3.5 Further kinetics studies of NCS

Preliminary kinetics experiments performed by Samanani *et al.*[16,17,23] and Luk *et al.*[24] were performed in phosphate buffer, which is known to catalyse a racemic Pictet-Spengler reaction so a high non-enzymatic background reaction was observed. More recent kinetics experiments were performed by Lichman *et al.*[28] in HEPES buffer, with a minimal background reaction observed. Contradicting previous experiments, co-operativity was not observed and at saturating levels of 4-HPAA (**6**) and varying levels of dopamine (**5**), values of k_{cat} and K_m were shown to be 24 s^{-1} and 22 mM, respectively, compared with values of 4.5 s^{-1} and 0.4 mM which were observed previously. NCS was,

therefore, demonstrated to have very low catalytic efficiency with k_{cat}/K_m = 1.1 s^{-1} mM^{-1} whereas the median value for all enzymes is 125 s^{-1} mM^{-1}.[34] High substrate loading is therefore required for highly productive NCS reactions.

Further attempts to understand the kinetics of the NCS mechanism have proved challenging due to the many intermediates involved.[31] It is possible that several of the reaction steps could occur either inside or outside the active site; intermediates can diffuse in and out of the active site. Understanding is complicated by the possible isomers and protonation states of reaction intermediates. The racemic, background reaction observed has also complicated studies as saturating substrate concentrations cannot be used.[31]

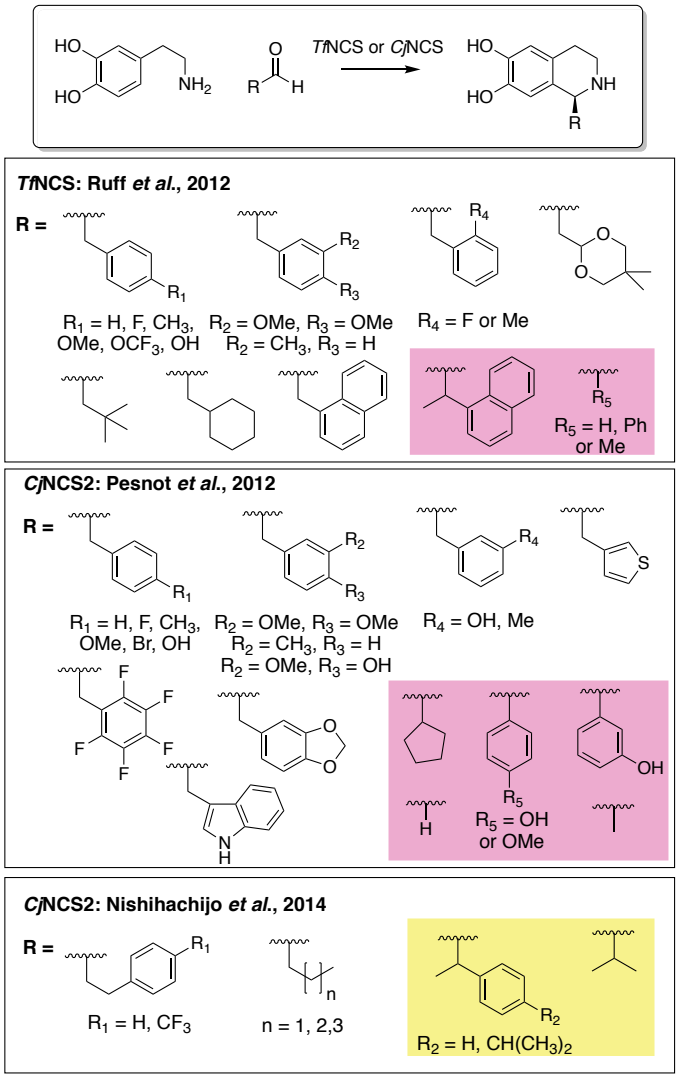
1.4 Substrate scope of norcoclaurine synthases

1.4.1 Aldehyde substrate scope

Expanding the substrate scope of NCS beyond that observed naturally is valuable for developing stereoselective syntheses of THIQs. The first attempts to expand the NCS substrate scope beyond dopamine (**5**) and benzylic aldehydes were performed by Ruff *et al*.[21] and Pesnot *et al*.[22] using two different NCSs, Δ29*Tf*NCS and *Cj*NCS2, respectively. Studies of Δ29*Tf*NCS showed a versatile aldehyde substrate scope when using dopamine as the amine donor. A wide range of aldehydes, from those with bicyclic heterocyclic and linear aliphatic substituents were successfully turned over and THIQ products generated in high yields (most >50% after 3 h), although enantiopurities of the products were not determined. Alpha-substituted aldehydes, benzaldehyde and 1-α-methyl-napthylacetaldehyde and small aldehydes, acetaldehyde and formaldehyde were not accepted.[21] This led to the assumption that alpha-substituted aldehydes, generally are not accepted as substrates by *Tf*NCS, rationalised as being due to the α-carbon being deeply buried in the enzyme active site, leading to a lack of conformation freedom of the iminium ion intermediate.[28] Subsequently, low activities of *Tf*NCS with

benzaldehyde and cyclohexanecarboxaldehyde have been observed, however, product formation and enantioselectivity have not been verified.[31] The Δ29*Tf*NCS variant, L76A has been shown to modulate aldehyde acceptance. Compared with the wild-type enzyme, increased conversions were observed with both enantiomers of citronellal, and in particular were enhanced with the (*S*)-enantiomer and decreased conversions were observed with hexanal and 4-HPAA.[28]

Similar activities were observed with *Cj*NCS2 in studies performed by Pesnot *et al*., with a range of linear aliphatic and benzylic aldehydes being well-tolerated as substrates[22] and investigations demonstrated that the THIQ products were generated in high e.e.s (>95%)[22,35] and low activities were also observed with alpha-methyl substituted aldehydes.[35] The investigations by Nishihachijo *et al*. also revealed the acceptance of linear, aliphatic aldehydes by *Cj*NCS2.[35] The observed aldehyde substrate scopes of both variants are given in Scheme 1.8.



Scheme 1.8: The observed aldehyde substrate scopes of *Tf*NCS and *Cj*NCS2. A range of different C-1 substituted (*S*)-THIQs can be synthesised. Given in pink, are the aldehydes tested where no or trace conversions were observed. Given in yellow are aldehydes where very low conversions were observed.

1.4.2 Expansion of the carbonyl substrate scope

The carbonyl substrate scope of Δ29*Tf*NCS has recently been expanded towards ketones, leading to the formation of chiral 1,1-disubstituted and spirocyclic THIQs.[36,37] Chiral synthetic routes towards these THIQs is challenging and has involved the use of chiral formamidine auxiliaries and titanium-catalysed acylations.[38,39] A range of benzylic methyl ketones and cyclohexanone analogues were accepted, and the activity shown to be modulated by a range of single point variants. Improved conversions were

observed with hydrophobic substrates and variants A79I and A79F, likely due to increased hydrophobic interactions. A diketone was also shown to be accepted, with different variants giving different ratios of mono- and disubstituted Pictet-Spengler product formed i.e. reaction at one or both of the carbonyl groups (Figure 1.3). Variants A79I and A79F favoured mono Pictet-Spengler product whereas F80L favoured diproduct formation.[37] Assays involving *Cj*NCS2 have shown that the enzyme does not accept ketones as substrates, and the lack of activity is still not understood. The key active site residues in *Tf*NCS and *Cj*NCS2 are mostly conserved, other than at position 79 (A in *Tf*NCS and I in *Cj*NCS2) and there is an extra alanine in this loop (residues 76-80) in *Cj*NCS2.[22]

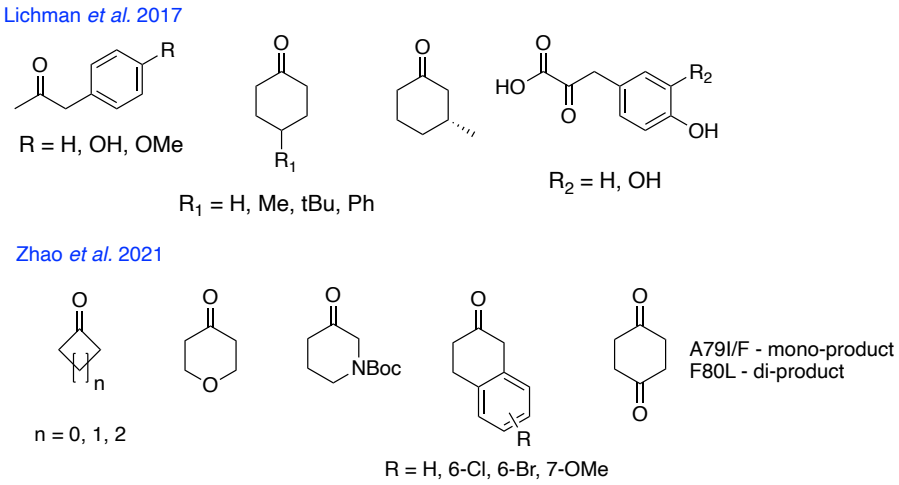


Figure 1.3: Non-aldehyde carbonyl substrate scope of *Tf*NCS explored to date.

Several other NCS variants have also been explored. Mutations of M97 were investigated as the sidechain is thought to occupy space near to the iminium ion intermediate. It was hoped that reducing steric bulk and flexibility of the amino acid side chain (by mutation to leucine or valine) could result in increased conformational freedom of the iminium ion intermediate and thus improve substrate conversions. Unfortunately, these mutations resulted in a decrease in activity with hexanal and phenylacetaldehyde as aldehyde substrates.[36] The use of DMSO as the reaction co-solvent also resulted in a five-fold increase in activity compared with using acetonitrile, and it was hypothesised that this was due to altering the reaction equilibrium to favour

iminium ion formation and it is also possible that this may have simply aided substrate solubility.

Several alpha-keto acids have also shown to be accepted as substrates for *Tf*NCS and *Cj*NCS2. 3,4-Dihydroxyphenylpyruvate (**13**) was initially hypothesised to be the natural carbonyl substrate, followed by decarboxylation to give (*S*)-norlaudanosoline (**15**).[12] 4-Hydroxyphenylpyruvate has also been reported to be accepted as a substrate, however enzymatic reactions were performed in phosphate buffer which is known to catalyse a racemic Pictet-Spengler reaction.[15] Activity has also been observed for the *Tf*NCS catalysed reaction between dopamine (**5**) and α-ketoglutaric acid, although the product formed was not isolated.[31] The carbonyl substrate scope of *Tf*NCS beyond aldehydes is given in Figure 1.3.

1.4.3 Amine substrate scope

A narrow amine substrate scope has been observed for both *Tf*NCS and *Cj*NCS2. The *meta*-hydroxyl group of dopamine (**5**) is essential for catalytic activity and removal or methoxy- substitutions of the *meta*-hydroxyl group are currently not tolerated. These modifications are tolerated on the *para*-hydroxyl group. Minor substitutions on the alkyl region of dopamine are also tolerated to date, with metaraminol and norepinephrine accepted as substrates, although few analogues have been screened.

1.5 Phylogenetic analysis of norcoclaurine synthase

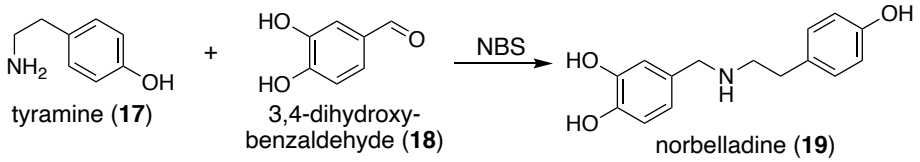
1.5.1 Proteins related to norcoclaurine synthase

NCS has the most significant sequence similarity with the PR10 and Bet v 1 family of proteins. PR10 proteins are a sub-class of pathogenesis-related proteins which are upregulated by plants under stress. Despite a wealth of structural insights and the ubiquitous nature of the PR10 proteins, the *in planta* role is largely unknown.[40] Bet v 1 proteins are a subclass of the PR10 proteins with a high sequence similarity to NCS

(50 - 60%) and are widely studied due to the allergic response in around 25% of humans.[41]

Enzymatic activity in proteins with a PR10 fold is uncommon. There are few known examples beyond NCS. A fungal enzyme involved in polyketide tetracenomycin aromatase/cyclase, biosynthesis has low sequence similarity to PR10 proteins, but the topology is remarkably similar.[42] Two other enzymes involved in BIA-biosynthesis, neopinone isomerase (NISO) and thebaine synthase (THS) are also classified as PR10/Bet v 1-like proteins.[43]

Recently, an enzyme involved in Amaryllidaceae alkaloid biosynthesis has been identified with high sequence identity to NCS (38% identity to *Tf*NCS) considering that other NCS homologues also have this level of similarity to each other. The enzyme in question, norbelladine synthase (NBS), has been isolated from *Narcissus pseudonarcissus* and was shown to catalyse the condensation of tyramine (**17**) and 3,4-dihydroxybenzaldehyde (**18**) to give norbelladine (**19**) as shown in Scheme 1.9.[44]



Scheme 1.9: The biosynthetic role of norbelladine synthase (NBS). The enzyme synthesises norbelladine, the first committed intermediate towards the Amaryllidaceae alkaloids.

The mechanism of NBS is thought to proceed via the condensation of the two substrates to form an iminium ion intermediate. No additional co-factors are required for a successful enzymatic reaction. However, 30 equivalents of aldehyde with respect to the amine are required in *in vitro* reactions, suggesting that the aldehyde (or hydrate) is acting as the reducing agent by providing a hydride source. It is unknown whether this reduction step is catalysed by another enzyme *in planta*.[44]

NCSs are unique compared with other PR10 proteins due the presence of *N*- and *C*-terminal extensions, which are suspected to be involved in subcellular localisation. NBS lacks *N*-terminal residues 1-41 compared with NCS and these residues have been shown to be non-essential for the catalytic activity of NCS.

Recent phylogenetic analysis of *N. nucifera*, the sacred lotus, has identified two different subfamilies of NCS: NCSI and NCSII. NCSII genes are found in all plants whereas NCSI genes are only found in BIA-producing plants. It is hypothesised that NCSI genes have evolved from the duplication of NCSII genes. No recombinant proteins encoded for by NCSII genes have been isolated and characterised, so it is unknown whether they have any NCS- like activity. Some non-BIA producing plants have been shown to possess enzymes with similar activity to NCS, however, this has not been fully investigated.[45,46]

1.5.2 The PR10 fold

There have been structural investigations of numerous proteins within the PR10 family, and a common protein fold has been noted. The fold consists of a seven-stranded anti-parallel β-sheet wrapped around a *C*-terminal α-helix, giving a 'baseball glove – like' framework. There are two additional alpha-helices located between the sixth and seventh *N*-terminal beta-sheet strands. A large (*ca.* 3900 $Å^3$), internal hydrophobic cavity is also present, which is thought to be functionally relevant to many PR10 proteins and has been shown to be a ligand binding site.[40] The common PR10 fold is given in Figure 1.4.

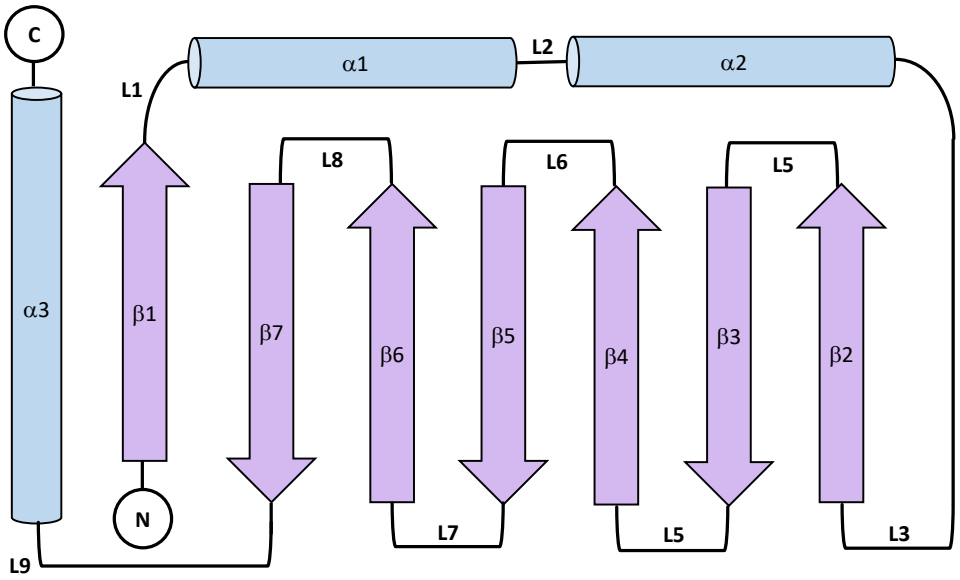


Figure 1.4: **The common PR10 fold.** The connectivity of the different secondary structure elements is given. β-sheets (arrows) and α-helices (cylinders) are connected by 9 loops (denoted L1-9). The *N*- and *C*- termini are given by circles. Adapted from Figure 2A of Fernandes *et al*.[40] The tertiary fold is demonstrated by the fold of *Tf*NCS given in Figure 1.2a, in the same colour scheme.

1.5.3 Native NCS enzymes

To date, several NCSs capable of catalysing a Pictet-Spengler reaction to give norcoclaurine have been identified and isolated. Those most extensively explored are those isolated from *Coptis japonica* and *Thalictrum flavum*, as previously mentioned and other variants in *Argemona mexicana*, *Corydalis Saxicola* and *Papaver braceatum* have been shown to be highly productive. However, the substrates scopes have not been investigated beyond the natural reaction.[20] In all cases, (*S*)-norcoclaurine has been isolated from these plants and key active site residues in the NCSs are conserved, suggesting that all act by a similar mechanism.

Nelumbo nucifera (sacred lotus) is an aquatic, BIA-producing plant found throughout Eastern Asia. The plant has been used in traditional medicine for centuries to treat cancers, bacterial infections and other illnesses.[47] Identifying and isolating the active components in the sacred lotus is therefore of scientific interest. *N. nucifera* is unique

amongst BIA-producing plants due to the apparent isolation of (*R*)-norcoclaurine from the seed embryo.[48,49] (*R*)-Coclaurine and (*S*)-norcoclaurine (**7**) have also been isolated from the leaves of the plant by solvent extraction and chromatography and identified by NMR, optical rotations and melting point data.[50] The biosynthesis of (*R*)-norcoclaurine is unknown, but it could arise from; the *in vivo* epimerisation of (*S*)-norcoclaurine (**7**), a non-enzymatic, racemic Pictet-Spengler reaction with selective transport and accumulation of each enantiomer in different parts of the plant or condensation by an (*R*)-selective Pictet-Spenglerase. Seven different genes encoding for NCS have been identified in the lotus and characterised into two different families: NCSI and NCSII. The genes of the NCSI subfamily are found in plants which express BIAs whereas the genes encoding for the NCSII subfamily are found in all plants.[46] One of the NCSs noted was found to be of the NCSII subfamily and of the six in subfamily NCSI, two were found to be pseudogenes (i.e., non-functional copies of others) and so there are four different NCSs which may be responsible for the production of (*R*)-norcoclaurine. The enzyme designated *Nn*NCS7 was found to be most significantly correlated with alkaloid content in the leaf in most cultivars. The paper gave no mention of alkaloid content in the seed embryo or which NCSs are most highly expressed there. Interestingly, unlike previously investigated NCSs such as *Tf*NCS and *Cj*NCS2, the lotus NCSs do not contain any signal peptides and so are likely to be localised in the cytoplasm rather than in vesicular compartments.[51]

1.6 Tetrahydroisoquinolines

1.6.1 Medicinal significance of tetrahydroisoquinolines

The THIQ moiety is found in many pharmacologically relevant molecules which act on a range of different receptors. It is therefore known as a privileged scaffold, so there is an interest in developing a range of analogues for drug discovery efforts. A range of useful biological properties have been observed for drug candidates containing a THIQ scaffold, including anti-tubercular[52], anti-cancer[53] and anti-HIV[50]. The family of

molecules also includes tetrabenazine, which is used for hyperkinetic movement disorders and was the first drug marketed to treat Huntington's disorder and Tourette's syndrome.[54] The THIQ, solifenacin is a widely prescribed antimuscarinic.[55] Developing facile, economically-feasible synthetic routes towards THIQ is therefore highly valuable.

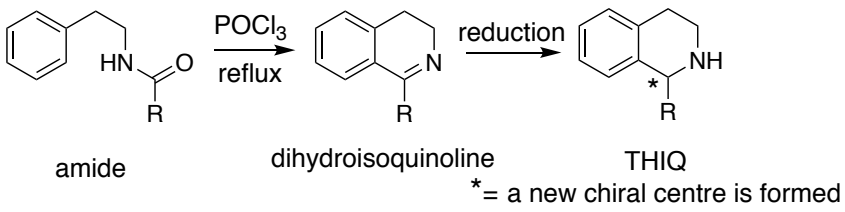
When assessing the pharmacological properties of chiral molecules, it is important to determine the activities of both enantiomers as they may vary in terms of toxicity, potency and selectivity of receptors and rate of metabolism.[56] For example, (*S*)-norcoclaurine (**7**) has a higher inhibitory effect on epinephrine-induced platelet aggregation than the corresponding (*R*)-enantiomer.[57] As many biologically-relevant THIQs contain one or more chiral centres, it is also important to be able to isolate each isomer so that the activities of each can be assessed.

1.6.2 Routes towards racemic THIQs

There are several well-established synthetic routes towards racemic formulations of THIQs. The two most employed strategies are via a Pictet-Spengler reaction (discussed in Section 1.1.2) or a Bischler-Napieralski reaction. The Pictet-Spengler reaction is a single-step route involving the condensation between a phenethylamine and an aldehyde or ketone. Most often, the reaction is acid-catalysed, however, more recently a phosphate-catalysed reaction has been realised, leading to the synthesis of a range of C-1 substituted THIQs under mild conditions.[15] The reaction proceeds via a two-step mechanism, whereby one phosphate bridges the iminium ion and the aromatic ring, both aiding the cyclisation reaction and deprotonating the transition state to regain aromaticity in the product. Another phosphate deprotonates the *meta*-hydroxyl group to increase the ring nucleophilicity and also reprotonates this hydroxyl during the cyclisation reaction.[58] The phosphate-mediated reaction has also avoided the use of high temperature reactions when using ketones as substrates, conditions which also usually preclude the use of oxidatively-sensitive substrates such as catechols.[59] The

main limitations of the racemic, chemical Pictet-Spengler reaction are the challenges of generating single regioisomer products and the requirement for a *meta*-hydroxyl group on the phenethylamine. Nucleophilic attack of the aryl ring π-system towards the iminium ion carbon can also proceed from the *ortho* or *para* position, so two regioisomers of product are generated. The ratio of the two regioisomers generated is dependent on the pH of the reaction media[60] and the carbonyl substrate used.[15] Separation of the two regioisomers is challenging and can require preparative HPLC so is not feasible on larger scales. Interestingly, when performing the reaction between metaraminol and 2-bromobenzaldehyde, stereochemical control at the C-1 position was observed, with the (1*S*,3*S*,4*R*) product generated in >97% diastereomeric excess.[61]

The Bischler-Napieralski reaction is an intramolecular electrophilic aromatic substitution reaction of an amide, leading to the formation of a dihydroisoquinoline.[62] This can be followed by a reduction to give the THIQ scaffold (Scheme 1.10). The Bischler-Napieralski reaction requires a strongly dehydrating Lewis acid, most often phosphorus oxychloride ($POCl_3$). Although generating a multi-step strategy is required to generate the THIQ scaffold, regioselective control is achieved and only the *para*-isomer of product is normally generated.



Scheme 1.10: The general synthetic route to generate the THIQ scaffold when a Bischler-Napieralski reaction is involved.

1.6.3 Chemical routes towards single enantiomer tetrahydroisoquinolines

The synthesis of the THIQ scaffold with a Pictet-Spengler or Bischler-Napieralski reaction leads to the generation of a new chiral centre at the C-1 position. With both

synthetic strategies, stereochemical control can be achieved. Asymmetric Pictet-Spengler reactions can be performed using a chiral Brønsted acid catalyst[63, 64] or a chiral auxillary.[65] If a chiral hydrogenation catalyst (usually rhodium, iridium or ruthenium-based) is used for the reduction step, single regioisomer and enantiomer THIQs can be generated via a Bischler-Napieralski route. Although many synthetic routes towards single isomer THIQs have been reported, there are no general routes that can be applied to all substrates.[66]

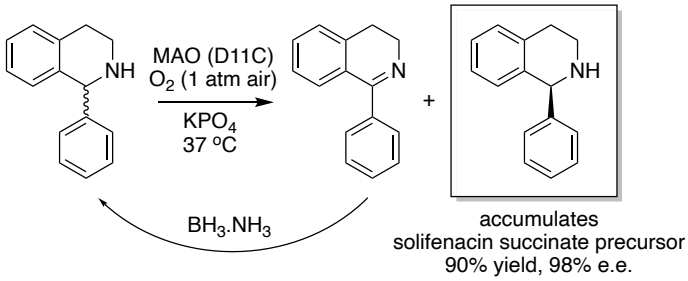
1.6.4 Biocatalytic routes towards single isomer tetrahydroisoquinolines

Biocatalysis is being increasingly employed for single enantiomer syntheses, with enantiopurity generated via either the kinetic resolution of a racemic formulation or stereochemical control when a new chiral centre is generated in a reaction. Enzymes are advantageous over traditional synthetic methods as reactions are performed under mild conditions and they often exhibit exquisite stereoselectivity. However, high reaction selectivity can go 'hand in hand' with narrow substrate scopes, poor versatility and the discovery of novel enzymes is challenging and requires significant research efforts. However, as the repertoire of known enzymes and their sequences increases, these challenges are diminishing, and a large range of chemical transformations can be performed using recombinantly expressed enzymes.

As discussed in Section 1.4, *Tf*NCS is a versatile catalyst and can generate a wide range of C-1 substituted, (*S*)-THIQs in high yields and e.e.s. Many carbonyl substrates are accepted, however the *meta*-hydroxyl group on the phenethylamine is required for a successful reaction to occur and only the (*S*)-enantiomer of product can be generated to date. There are also other biocatalytic routes towards THIQs.

Monoamine oxidases (MAOs) catalyse the oxidation of an amine to an imine. A kinetic resolution of the amine is performed, with one stereoisomer of the amine preferentially

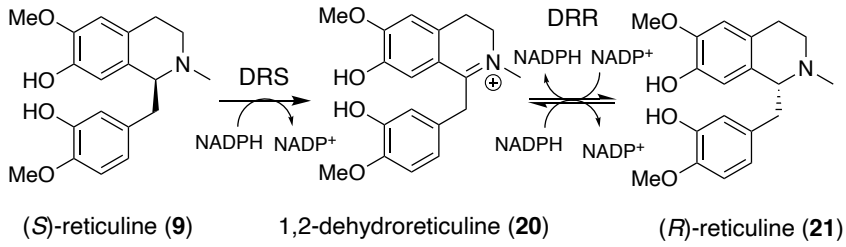
oxidised. The imine can then be reduced *in situ* to regenerate the racemic amine and over time, this leads to the accumulation of a single enantiomer of amine, of opposing stereochemistry to the amine oxidised by the MAO. One MAO has been developed by Ghislieri *et al.*[67] which accepted racemic THIQs. The enzyme was developed using a combination of structure-guided engineering and high-throughput screening. The novel enzyme was used to generate (1*S*)-1-phenyl-1,2,3,4-tetrahydroisoquinoline, the main synthetic precursor of solifenacin succinate in 90% yield and 98% e.e. (Scheme 1.11). Although a highly productive reaction was observed, there is a synthetic requirement to generate the racemic THIQ precursor and the substrate scope with other THIQs has not yet been explored. Moreover, as an oxidising enzyme, oxidatively sensitive substrates cannot be used.



Scheme 1.11: The generation of (1S)-1-phenyl-1,2,3,4-tetrahydroisoquinoline using an engineered MAO. The *in situ* reduction of the imine formed results in the accumulation of the THIQ enantiomer that is not oxidised by the MAO.

An alternative approach is the reduction of imines. Instead of using a chiral hydrogenation catalyst, the stereoselective reduction of a dihydroisoquinoline to a THIQ can be performed using an imine reductase. A number of imine reductases have been explored and are capable of generating THIQ products in high e.e.s.[68–71] The reported substrate scopes are, however, limited to either no or methoxy substituents on the phenyl ring; i.e. no oxidatively sensitive hydroxyl groups are tolerated, and there is the synthetic requirement to generate the dihydroisoquinoline substrate via a Bischler-Napieralski reaction.

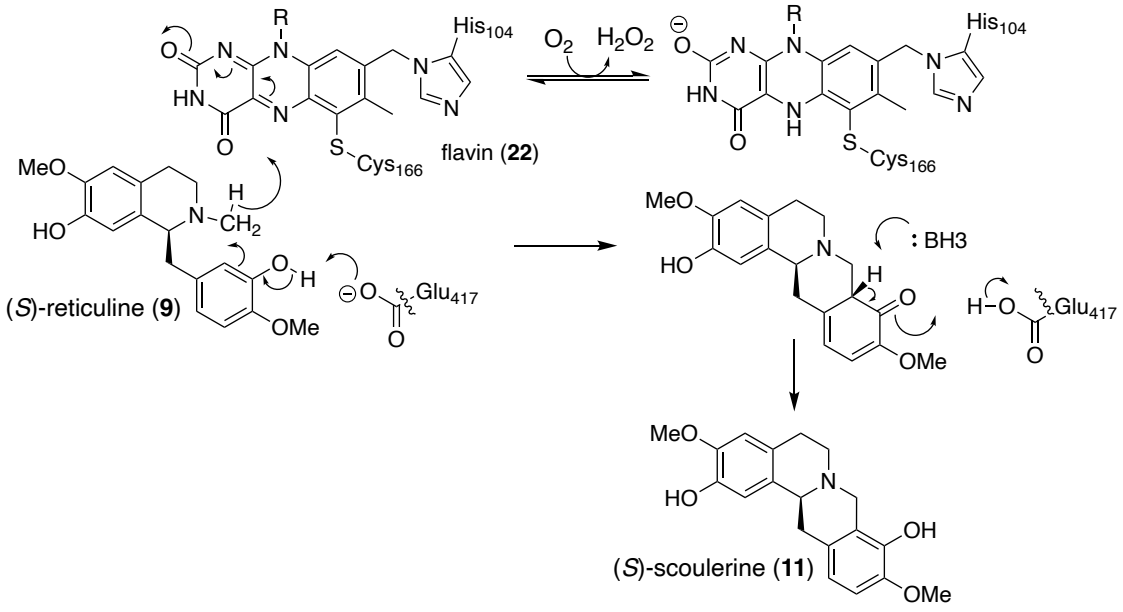
Two other enzymes have potential for generating single-enantiomer THIQs, although both have not been widely explored. Firstly, the reticuline epimerase enzyme (REPI or STORR) which is responsible for the inversion of stereochemistry at the C-1 position of (*S*)-reticuline (**9**) to (*R*)-reticuline (**21**).[7,72] REPI was the final enzyme to be isolated in the biosynthetic pathway towards opiate alkaloids and has since led to the *in vivo* production of morphine in yeast.[73] The enzyme is a cytochrome P450 fusion protein with two domains; 1,2-dehydroreticuline synthase (DRS) oxidises (*S*)-reticuline (**9**) towards, 1,2-dehydroreticuline (**20**) and 1,2-dehydroreticuline reductase (DRR) reduces the iminium ion to give (*R*)-reticuline (**21**). A range of *O*- and *N*-methylated analogues of **9** have been investigated as substrates, for REPI, DRS and DRR. The *N*-methyl group is essential for catalytic activity and various methylation patterns on the *O*-heteroatoms are tolerated. No structural studies of the enzyme have been performed and the mechanism of action is unknown. This information would be valuable to develop variants via structure-guided mutations to lead to an enzyme capable of generating the (*R*)-analogues of THIQs synthesised by an NCS-mediated synthesis. The reaction pathway of REPI is given in Scheme 1.12.



Scheme 1.12: The reaction pathway of the reticuline epimerase enzyme.

Another enzyme that may be valuable in generating single enantiomer THIQs is the berberine bridge enzyme (BBE). The enzyme catalyses an intramolecular carbon-carbon bond forming reaction between the *N*-methyl group of (*S*)-reticuline (**9**) and the aryl ring, forming a new six-membered ring to give (*S*)-scoulerine (**11,** Scheme 1.1). This gives the tetracyclic berberine scaffold, the formation of which is essential in the biosynthetic pathways towards benzophenanthridine and protoberberine alkaloids.[74]

Structural evidence for the mechanism has been gained from a crystal structure of the enzyme with (*S*)-reticuline (**9**) bound in the enzyme active site.[74,75] The enzyme requires a flavin (**22**) as a co-factor for substrate turnover and the proposed mechanism of action is given in Scheme 1.13.



Scheme 1.13: The proposed reaction mechanism of the BBE. Interestingly, the reaction is proposed to proceed via the loss of a hydride leaving group which is incorporated into the flavin co-factor. Adapted from Scheme 1 of Winker *et al*.[74] The co-factor is regenerated by reaction with molecular oxygen, giving hydrogen peroxide. Co-factor re-oxidation was measured using a stopped-flow instrument.

The reaction is initiated by the deprotonation of the *meta*-hydroxyl of **9** by E417, leading to a carbon-carbon bond forming reaction and the incorporation of the hydride leaving group into the flavin co-factor (**22**). Rearomatisation gives the final product, (*S*)-scoulerine (**11**) and the flavin co-factor (**22**) is regenerated by oxidation with molecular oxygen. The substrate scope of heterologously expressed BBE (in *Pichia pastoris*) from *Eschscholzia californica* has been explored. The *N*-methyl group is essential for catalytic activity and various methylation patterns on the aryl rings are tolerated and activity is still retained while these hydroxy and methoxy groups are removed. The

enzyme is selective for substrates with (*S*) stereochemistry at the C-1 position, a feature which has been exploited to give a range of (*R*)-benzylisoquinoline derivatives.[76] The regioselectivity of the ring closure step can also be inverted when fluorinated substrates are used.[77]

1.7 Other Pictet-Spenglerases

NCS is not the only Pictet-Spenglerase enzyme, many others have been identified as performing a key biosynthetic role in alkaloid production. Pictet-Spenglerases are lyases (EC 4) but there is no dedicated sub-class. This section discusses other known Pictet-Spenglerases, their biosynthetic roles and their biocatalytic applications. The natural substrate scopes of the known Pictet-Spenglerases, with the products generated are given in Figure 1.5.

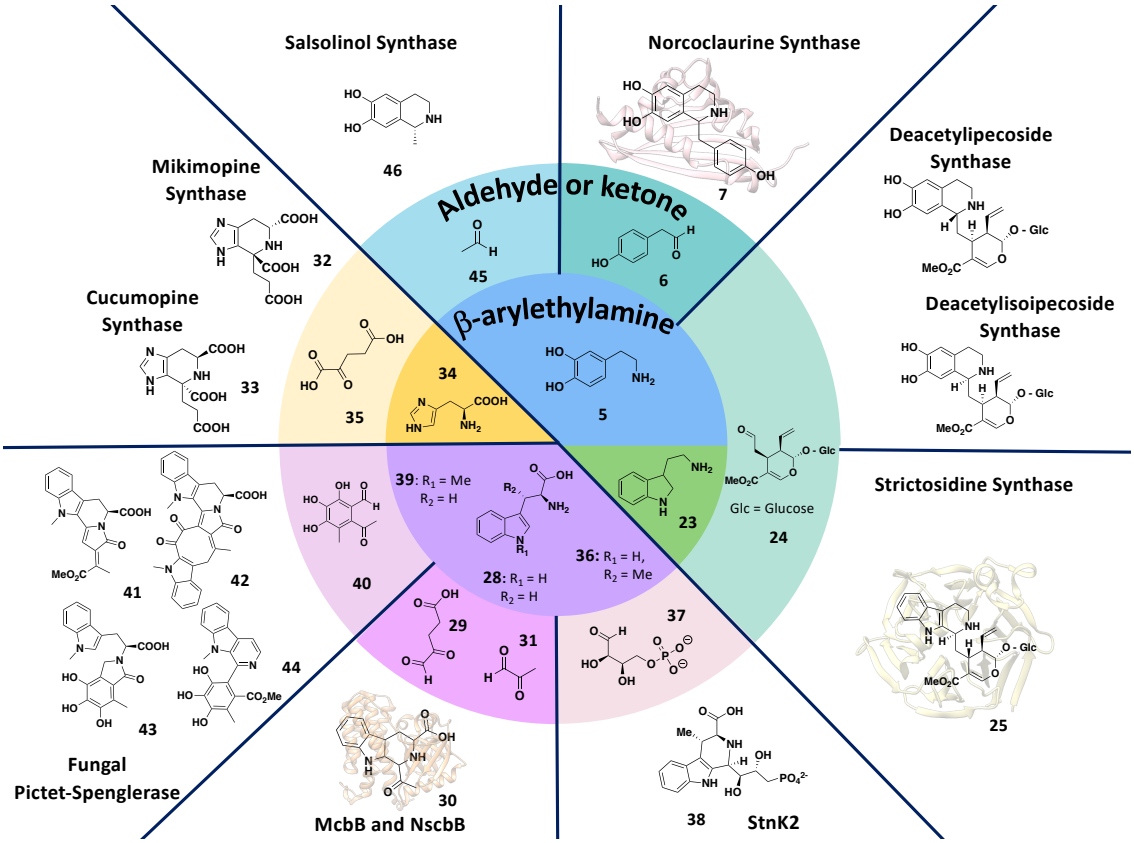


Figure 1.5: The natural substrate scopes of known Pictet-Spenglerases. Each segment corresponds to an enzyme, which condenses a β-arylethylamine (inner circle) and a carbonyl-containing substrate (middle circle) to give the product (outer). Protein structures, where known, are represented. Figure has been adapted from Figure 1 of Roddan *et al*.[19]

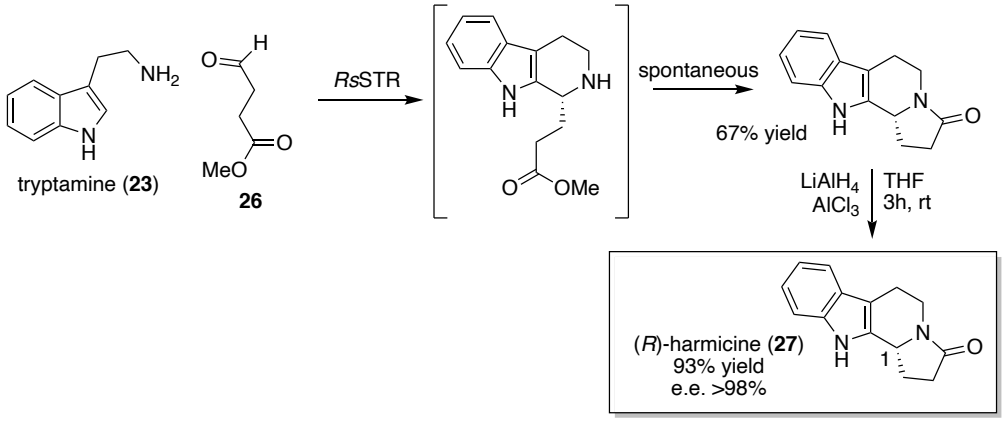
1.7.1 Strictosidine synthase

Strictosidine synthase (STR) catalyses the Pictet-Spengler reaction between tryptamine (**23**) and secologanin (**24**) to form 3-α(*S*)-strictosidine (**25**), the biosynthetic precursor to monoterpenoid indole alkaloids.[78] There are over 2,000 different monoterpenoid indole alkaloids (MIAs), many of which have important medicinal activities, such as quinine (anti-malarial), camptothecin (anti-tumour agent) and ajmaline (anti-arrhythmic).[79] Two STRs have been most extensively characterised, isolated from *Rauvolfia serpentia* (*Rs*STR) and *Catharanthus roseus* (*Cr*STR).[80,81]

The mechanism of action was elucidated by Maresh *et al.*[82] in 2008. A crystal structure of recombinant *Rs*STR co-crystallised in the presence of the two natural substrates, secologanin (**24**) and tryptamine (**23**) was solved. The structure of STR is unusual, being the first example of a β-propeller protein found in the plant kingdom.[81] Mutagenesis identified the key active site residues as E309, Y151 and H307. Kinetic isotope effects and pH dependence of the reaction suggested that formation of the iminium ion intermediate is acid-catalysed and that the final deprotonation step is base-catalysed. *Ab initio* calculations have indicated that the reaction mechanism does not go via a spiroindolenine intermediate (similar to that initially suggested for the NCS reaction, Scheme 1.5).[82]

The amine donor substrate scope is not limited to tryptamine and screens have been performed with wild-type and mutated STRs.[83–85] Analogues with hydroxyl and methoxy groups at C-5 and C-6 respectively of tryptamine (**23**) were accepted.[86] Activity was also retained when the tryptamine benzene ring was substituted for a pyridine moiety or if the pyrrole ring of **23** was exchanged for a furan ring.[87] The aldehyde substrate scope has also been altered using active site mutants.[85,86,88]

Despite these successes, recombinant expression and isolation of STRs has proven challenging. A recent paper by Pressnitz *et al*. improved the expression and activity of clarified cell lysates 100-fold via optimisation of the expression protocol and removal of signal peptides.[89] Interestingly, a variety of non-natural, aliphatic aldehydes have been accepted by four different STRs to give single enantiomer products with (*R*)-stereochemistry at the C-1 position. This led to the enantioselective chemoenzymatic synthesis of (*R*)-harmicine (via the condensation of **23** and **26**), in an analogous method to the synthesis of trolline-derivatives using *Tf*NCS (Scheme 1.14).[89] Rationale for the inversion of configuration was gained from a combination of molecular dynamics studies and a crystal structure of *Op*STR (isolated from *Ophiorrhiza pumila*) with a non-productive reaction intermediate analogue bound in the active site, using an analogous design to the mimic used to probe the NCS mechanism.[30,90] The structure shows the substrates binding in an inverted manner compared to reactions performed with the natural substrates.



Scheme 1.14: Two-step, chemoenzymatic synthesis of (*R*)-harmicine

STR has also been employed in the stereoselective synthesis of *N*-substituted strictosidine derivatives as novel topoisomerase inhibitors and other alkaloids via chemoenzymatic cascades.[84,91]

1.7.2 Deacetylipecoside synthase

Two different Pictet-Spenglerases have been identified in *Alangium lamarckii* with the ability to condense dopamine (**5**) and secologanin (**24**). Deacetylipecoside synthase (DIS) forms the (*R*)-enantiomer at C-1, while deacetylisoipecoside (DIIS) forms the (*S*)-enantiomer (Figure 1.5). Both undergo spontaneous lactamisation followed by subsequent enzymatic modifications to give alangiside and isoalangiside-type glucosides respectively. DIS has been successfully isolated and purified from *A. lamarckii* and found to be 30 kDa in molecular weight, however, DIIS was found to be too labile for purification.[92]

1.7.3 Bacterial Pictet-Spenglerases

Five different bacterial Pictet-Spenglerases have been identified; McbB, NscbB, mikimopine synthase, cucumopine synthase and the SfmC module of the non-ribosomal peptide synthetase (NRPS). The enzyme McbB has been found to perform the Pictet-Spengler reaction between L-tryptophan (**28**) and oxaloacetaldehyde (**29**) to give **30** (Figure 1.5). Subsequent oxidation and decarboxylation steps give a β-carboline scaffold, found in many pharmacologically active molecules including benzodiazepine inverse agonists.[93,94] The enzyme has been shown to accept the non-natural substrates, 5-methyl-DL-tryptophan, 7-methyl-DL-tryptophan and smaller aldehydes such as methylglyoxal, formaldehyde, acetaldehyde, propanal and isobutyraldehyde instead of oxaloacetaldehyde. The co-crystallised structure of McbB with L-tryptophan was obtained by Mori *et al*. in 2015. The active site is formed by a homodimerisation where each monomer adopts a slightly different conformation. The catalytically important active site residues have been determined and site-directed mutagenesis resulted in the formation of various active mutants. Mutations of two bulky residues, H87 and R72, located at the entrance of the active site, to alanine lead to the acceptance of the unnatural aldehyde, phenylglyoxal.[94]

NscbB (identified in *Nocardiopsis synnemataformans*, derived from a kidney transplant patient) catalyses the Pictet-Spengler reaction between L-tryptophan (**28**) and methylglyoxal (**31**) to give 1-acetyl-3-carboxy-β-carboline (**30**) i.e., the same reaction as with McbB (Figure 1.5). Both enzymes have high sequence identity (66%) and similarity (80%) with a conserved active site residue E97 suggesting that that both operate via similar reaction mechanisms. NscbB has *ca.* a 30 fold higher k_{cat}/K_M than McbB, but NscbB is less thermally stable.[95]

Mikimopine (**32**) and cucumopine (**33**) are opines, formed via the Pictet-Spengler condensation of histidine (**34**) and α-ketoglutaric acid (**35**) (Figure 1.5). Opines are found in plant tumours induced by the parasite, *Agrobacterium*. The T-DNA encoding for the enzymes responsible for opine biosynthesis are passed to the plant via horizontal gene transfer. Opines are then synthesised by the plant cells and provide carbon and nitrogen sources for the invading bacteria. There has been little characterisation of the two enzymes; however, the genes encoding mikimopine synthase have been isolated from *A. rhizogenes* and both enzymes have been expressed recombinantly in *E. coli* and enzymatic activity confirmed.[96–98] Interestingly, the carbonyl substrate is an α-ketoacid, whereas other Pictet-Spenglerases typically have an aldehyde as the natural substrate.

Stnk2 is a Pictet-Spenglerase involved in the biosynthesis of streptonigrin, an alkaloid antibiotic with antitumor activity and it has high sequence identity with McbB (41%). The enzyme performs a stereospecific reaction between (2*S*,3*S*)-β-methyl-tryptophan (**36**) and D-erythrose-4-phosphate (**37**), with the newly formed chiral centre in the (*R*)-configuration (**38**) (Figure 1.5). (*S*)-Stereochemistry at C-3 in tryptophan is essential for its reactivity. The aldehyde substrate scope is limited: methylglyoxal (**31**) and ethyl glyoxalate were not accepted. Interestingly, several fluoro-substituted L-tryptophan analogues were accepted, and improved affinity was observed with 5- and 6-fluoro-

(2*S*,3*S*)-β-methyl tryptophan compared to the natural substrate. This is, therefore, a promising strategy towards generating fluorinated analogues of streptonigrin.[99]

Finally, the NRPS SfmC module has been found to perform seven sequential reactions, including two Pictet-Spengler condensations in the biosynthesis of Saframycin C, an anti-tumour antibiotic with a THIQ scaffold.[100]

1.7.4 Fungal Pictet-Spenglerases

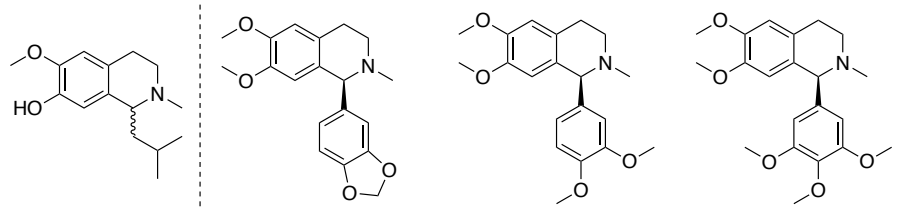
Comparative genetic analysis has been used to identify a silent Pictet-Spenglerase in the fish-derived fungi, *Chaetomonium globosum*. It was found that 1-methyl-L-tryptophan can upregulate the expression of the Pictet-Spenglerase which results in the condensation of 1-methyl-L-tryptophan (**39**) with the aldehyde, flavipin (**40**) (Figure 1.5). The product is then altered by other fungal enzymes to give a novel class of alkaloids the 'chaetoglines'(**41**-**44**) and pharmacological activities have been assessed.[101] Two have anti-bacterial activities and another acts as an inhibitor of acetylcholinesterase.[102]

1.7.5 Salsolinol synthase

Salsolinol has gained significant interest due to links with Parkinson's disease and alcoholism.[103] It is a THIQ alkaloid, formed by the condensation of dopamine (**5**) and acetaldehyde (**45**) (Figure 1.5). Higher levels of (*R*)-salsolinol have been found in the human brain, than the (*S*)-enantiomer based on chiral HPLC studies,[104,105] suggesting that salsolinol is formed enzymatically.[106] An enzyme, isolated from *Rattus norveticus,* has been overexpressed recombinantly in PC21 cells and expression correlated with an increased production of salsolinol. Chiral HPLC analysis of isolated salsolinol formed gave an e.e. of 20% (*R*-isomer).[107]

1.7.6 Putative Pictet-Spenglerases

THIQ alkaloids with an isobutyl group at C-1 (lophocereine (**47**)) have been isolated from *L. schotti,* a desert cactus (Figure 1.6).[108,109] No other naturally-occurring THIQs with an aliphatic group at the C-1 position have been identified. Feeding studies have suggested that both leucine and mevalonic acid are precursors to **47**. It is however known that leucine is not incorporated via mevalonic acid.[110,111] Although the enzymes present in *L. schottii* have not been identified, 3-methylbutanal was also incorporated during feeding studies suggesting that **47** may be formed by a Pictet-Spenglerase.[112,113]



Lophocereine (**47**) | (*S*)-(+)-Cryptostyline I (**48**) | (*S*)-(+)-Cryptostyline II (**49**) | (*S*)-(+)-Cryptostyline III (**50**)

Figure 1.6: THIQ alkaloids isolated from plants with unknown biosynthetic origins.

A range of enantiopure 1-phenyl-tetrahydroisoquinolines, known as the cryptostylines (**48**-**50**), have also been isolated from two species of orchid. In one species, *Cryptostylis erythroglossa*, only the (*R*)-enantiomer was isolated and in *C. fulva*, only the (*S*)-enantiomer was isolated (Figure 1.6).[114–116] The apparent enantiopurity of the isolated alkaloids suggests that the biosynthesis is enzymatic, likely via an epimerase or a Pictet-Spenglerase.

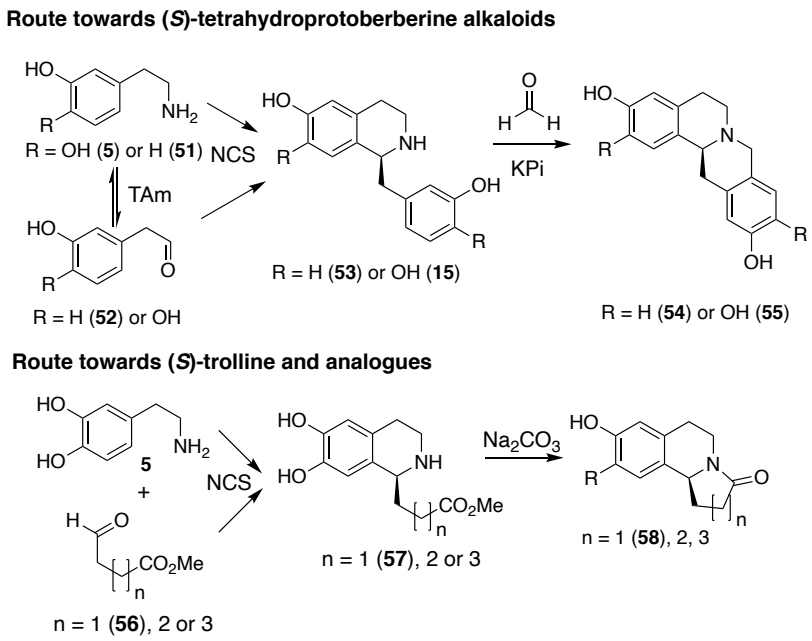
1.8 Synthetic biology routes towards THIQ syntheses

1.8.1 Biocatalytic cascades *in vitro*

NCS is a versatile biocatalyst, capable of generating a range of C-1 substituted THIQs beyond the natural product, (*S*)-norcoclaurine. Cascade processes involving NCS and other biocatalysts to generate THIQs are an attractive route to avoid the use of oxidatively-sensitive aldehydes as substrates and to generate further molecular complexity in the products generated. The aim is that this will lead to the facile,

environmentally benign production of a range of small molecules, containing the pharmaceutically relevant THIQ moiety. Both natural and unnatural THIQs can be generated in this manner, whilst avoiding the challenges of isolating compounds from the natural environment, where active components are often produced in low yields alongside other, structurally similar metabolites.

There are two examples of NCSs being employed in chemoenzymatic cascades, i.e., a cascade process involving both chemical and biocatalytic steps. (*S*)-Tetrahydroprotoberberine (THPB) alkaloids have been generated in a one-pot cascade involving three reaction steps. Firstly, a transamination reaction was performed with a phenethylamine (*meta*-tyramine (**51**) or dopamine (**5**)), to convert half of the amine into the analogous aldehydes, **52** and **18,** respectively. The amine and analogous aldehyde were then condensed via an NCS-mediated Pictet-Spengler reaction to give a 1-benzyl-THIQ (**53** and **15**). Finally, a phosphate-catalysed, racemic Pictet-Spengler reaction was performed between the 1-benzyl-THIQ and formaldehyde, giving the tetracyclic, tetrahydroprotoberberine scaffold (**54** and **55**). Products were generated in high isolated yield (up to 86%) and e.e. (99%) (Scheme 1.15).[117]



Scheme 1.15: Chemoenzymatic cascades involving *Tf*NCS. Routes were reported by Lichman *et al*.[117] and Zhao *et al*.[118]

Another chemoenzymatic cascade involving NCS was used to generate (*S*)-trolline (**58**) (an alkaloid with anti-viral properties) and analogues.[118,119] An NCS reaction was performed between a phenylethylamine (dopamine (**5**), metaraminol or fluorinated analogues) and a linear, aliphatic aldehyde with a terminal methyl ester (**56**), to give the THIQ scaffold (**57**). The addition of sodium carbonate led to the formation of a lactam ring, between the secondary amine of the THIQ and the methyl ester introduced from the aldehyde. The products generated in both chemoenzymatic cascades are given in Scheme 1.14.

A range of natural and non-natural BIAs have also been generated *in vitro*, starting from L-tyrosine and analogues. Tyrosinases, decarboxylates, transaminases and NCS were utilised leading to the generation of a range of products in high yield (35-99% conversion) and high e.e.s (often >90%). This route is advantageous over *in vivo* routes as steps can be introduced which are not biologically-compatible, leading to greater range of possible structures generated.[120]

1.8.2 Whole-cell approaches towards BIAs and related alkaloids

Another route towards the production of alkaloids, involves a whole-cell approach whereby all suitable enzymes for a biocatalytic transformation are co-expressed in the same host. This strategy is attractive for industrial scale syntheses, especially when numerous enzymes are required, as this route avoids the recombinant expression and isolation of each biocatalyst. The alkaloids generated are also highly complex and no cost effective total synthetic routes have been developed.[121] Due to the medicinal relevance of BIAs, there are a number of research efforts into developing routes for *in vivo* production.

Many of the challenges of using NCS in whole-cell BIA approaches are due the simultaneous production of racemic norcoclaurine, likely from a phosphate mediated background reaction. Early attempts towards BIA production bypassed the use of NCS and involved the production of (*R*,*S*)-norlaudanosoline towards (*R*,*S*)-reticuline and on to salutaridine, one of the key intermediates towards morphinan alkaloids, using *Saccharomyces cerevisiae* as the host.[122] Following this study, (*S*)-reticuline was produced in *Escherichia coli* starting from dopamine, then subsequent modifications towards magnoflorine and scoulerine (**11**) were performed using *S. cerevisae* as many of the enzymes involved in the later stages of BIA biosynthesis are not very active when expressed in bacteria.[123] (*S*)-Reticuline was produced to give a 11 mg L^{-1} culture volume, but this was later improved to give 54 mg L^{-1} when the NCS expression level was increased.[124] The pathway has also been extended to start from L-tyrosine (**4**). A fermentative *E. coli* based system has been used generate (*S*)-reticuline (**9**), with a final yield of 46 mg L^{-1}.[125]

The steps from thebaine towards morphine (**1**), codeine and other opiate BIAs were developed by Thodey *et al*. in 2014 in *S. cerevisiae* and optimisation of the platform resulted in high titres of products, up to 131 mg L^{-1} of culture volume.[126] The complete *de novo* production of (*S*)-reticuline in *S. cerevisiae* was completed in 2015,[127] however, the complete biosynthesis of opioids in yeast was only possible with the discovery of the final enzyme in the biosynthetic pathway, the reticuline epimerase, which converts (*S*)-reticuline (**9**) into the corresponding (*R*)-enantiomer.[7,72,128] Although low titres of product were achieved, the study provides a valuable starting point towards alternative routes for morphine (**1**) production, that avoid farming vast quantities of poppies.[128] Further development of an analogous system has led to the biosynthesis of noscapine and halogenated (*S*)-reticuline (**9**) derivatives in yeast using halogenated tyrosine analogoues.[129] Recently, the yeast production of **9** has been achieved at high levels (more than 3 g L^{-1}) by over 20 strain successive modifications on three metabolic

pathways, particularly by removing genes corresponding to redundant oxidoreductases which had thus far limited the production of the NCS substrate, 4-HPAA (**6**) which is highly oxidatively sensitive. Non-natural THIQs with aliphatic and benzylic substitutions at C-1 have also in 2019 been biosynthesised using a yeast strain engineered to maximise NCS activity.[130] *Cj*NCS2 has also been shown to be cytotoxic to *S. cerevisiae*, restricting the production of (*S*)-reticuline (**9**) *in vivo* and compartmentalization of NCS in the peroxisome helped to improve product titres.[131]

1.9 Project Aims

The overall aim of the PhD project was to further the understanding of the NCS reaction, and expand the accessible substrate scope of NCS, thus leading to the synthesis of novel alkaloids, either using the wild-type or variant constructs of *Tf*NCS.

When the project was started, significant efforts had been made to explore the natural substrate scope and mechanism of NCS. Most investigations had focused on *Tf*NCS, with mechanistic studies previously performed by Dr. Benjamin Lichman (Ward and Hailes groups), in collaboration with Dr. Altin Sula (Keep group), involving computational docking, crystallographic investigations and the generation of a library of single- and double-point mutants of *Tf*NCS. The substrate scope had also been explored and Dr. Daniel Méndez-Sánchez and Dr. Jianxiong Zhao (both Hailes group) were working on expanding the ketone substrate scope. Enzymatic cascade processes involving NCS, dehydrogenases, tyrosinases, regioselective *O*- and *N*-methyltransferases and/or transaminases were being developed by other members of Hailes and Ward groups (Dr. Daniel Méndez-Sánchez, Dr. Jianxiong Zhao, Dr. Yu Wang, Dr. Fabiana Subrizi and Benjamin Thair).

The aim of this project was to investigate the use of sterically crowded aldehydes, in particular, alpha-substituted aldehydes, as NCS substrates. The acceptance of the

benzaldehydes is of particular interest, to develop enantioselective routes towards novel (1*S*)-aryl-THIQs. Approaches to be explored include the screening of existing *Tf*NCS variants and the rational design and screening of novel variants, with the design process guided by crystallographic investigations of *Tf*NCS, with active site bound reaction intermediate mimics. Redesign efforts will also be focused towards switching the enantioselectivity of the NCS reaction to generate (*R*)-THIQs. Cascade processes will also be investigated to expand the range of (1*S*)-THIQs that can be generated via an NCS-mediated Pictet-Spengler reaction. Novel products generated can be isolated (by workup or preparative HPLC) and characterised by common analytical techniques, including NMR spectroscopy, mass spectrometry and chiral HPLC analysis to determine the structures and enantiopurities of the THIQs generated. Other avenues of investigation include structural studies of related Pictet-Spenglerases or NCSs isolated from other organisms.

Chapter 2: Novel Biocatalytic Routes Towards Tetrahydroisoquinoline Alkaloids

2.1 Introduction

2.1.1 Generating novel biocatalysts

Biocatalysts are enzymes or microbes that can be used in synthetic chemical transformations. As sustainability becomes increasingly relevant, biocatalysis is an attractive alternative to traditional synthetic methods as reactions are performed under mild conditions using aqueous media under ambient pressures and temperatures. However, one of the major limitations to the widespread applicability of biocatalysts in chemical syntheses is the lack of catalytic promiscuity. Enzymes are proteins that are evolved in nature to perform a specific catalytic role, often exhibiting high regio- and/or stereoselectivity so often the substrate scope of wild-type enzymes is limited. Enzymes may also exhibit poor stability and there are issues with reusability, although many of these shortcomings have been overcome by immobilization strategies. Therefore, to expand the possibilities of biocatalytic reactions, there are research efforts into the discovery of novel enzymes and into modifying known enzymes to alter their activities. As the number of known, sequenced enzymes with verified activities increases, new enzymes are rapidly discovered in newly sequenced genomes. Although the activities of each enzyme must be determined, this has expanded the portfolio of useful biocatalysts and enzymes are an increasingly viable alternative to traditional organo- and metallocatalysis.[132]

Protein engineering technologies have also led to the expansion of the substrate scopes of enzymes. There are two main approaches; rational design and directed evolution. Rational design involves designing and screening novel variants based upon protein structures or homology models. The results from substrate screens help to guide further, iterative rounds of mutagenesis and screening, leading to the generation of enzymes with improved activities. The directed evolution of enzymes was pioneered

by Frances Arnold, leading to the award of the Nobel Prize in 2018.[133] Directed evolution involves iterative rounds error-prone PCR (polymerase chain reaction). Each round generates a library of random mutants which are screened and at each stage, the most promising mutant is taken through to the next round. Both techniques have been used to develop enzymes with dramatically improved activities compared to the wild-type and enzymes have even been developed to form chemical bonds which are not known biologically.[134,135] There are also efforts into the computational, *de novo* design of enzymes of which there have been many successes, although development times are long.[136] However, as our understanding of how protein sequence relates to structure and function grows, *de novo* protein design may become an exciting area for the development of novel biocatalysts.

2.1.2 Studies of NCS

NCS is a useful biocatalyst, with the wild-type enzyme capable of generating numerous (1*S*)-THIQs. The products generated are relevant for drug discovery purposes and in many cases, there are no analogous chemical routes. To aid rationale design of the enzyme and increase the catalytic utility, various structural studies have been performed leading to significant mechanistic insights, as discussed in Section 1.3. However, there are still gaps in our understanding. For example, why the carbonyl substrate scope is so promiscuous i.e., the acceptance of ketones and aliphatic aldehydes and why activities vary between variants such as why ketones are accepted as substrates by *Tf*NCS but not *Cj*NCS2. The development of enzyme mutants has the benefit of both furthering understanding of the mechanistic process and potentially altering catalytic activity.

Several NCS variants have been developed by a previous PhD student, Benjamin Lichman.[31] Notable changes in activities were observed with a range of single point mutants of *Tf*NCS. Mutations E110D, E110Q and K122L reduced catalytic activity by

>90%, indicating that the chain length, position, and functionality at these two positions are all essential for a successful reaction. Reactions with F112L had 10-20% of the activity compared with the wild-type for all tested substrates, A79I increased activities (20-50%) with hexanal, and 4-HPAA and L76A showed increased activities with citronellal, particularly with the (*S*)-enantiomer, but not with hexanal. For reactions with ketones, M97F and M97V showed lower conversions compared with the wild-type, A79I and A79F showed improved activities. A range of double mutants, swapping the locations of key active site residues, were also developed with the aim of altering the regioselectivity of the enzyme towards the *ortho*-product. However, only F112K-K122F and K122F-F112K were investigated and shown to generate the *para*-product.[31]

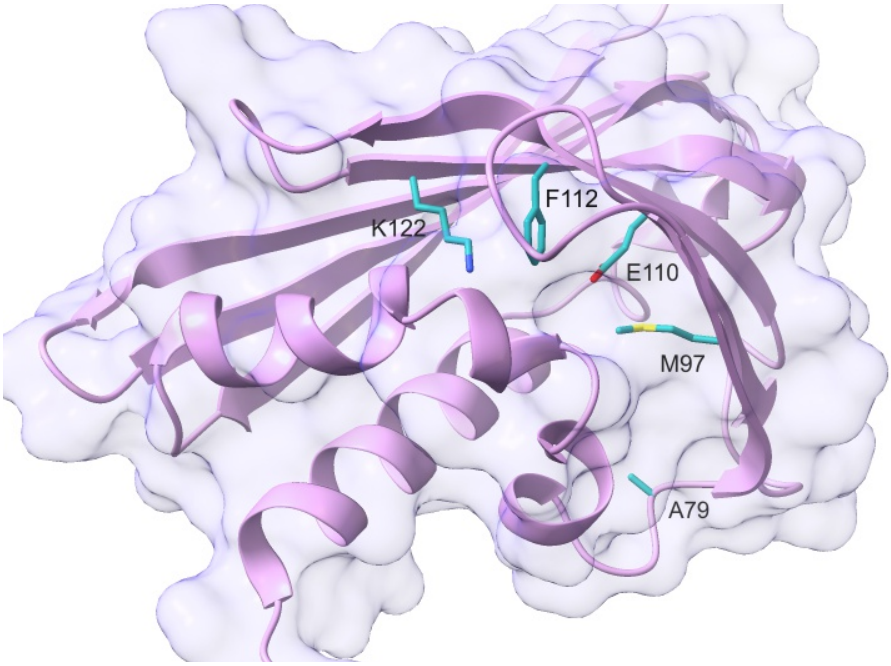


Figure 2.1: Structure of Δ29*Tf*NCS with key active site residues altered in previous studies shown in blue. The crystal structure shown was gained by Illari *et al*.[27] (PDB: 5N8Q) and the figure was generated using USCF Chimera.

Therefore, there is potential to further alter NCS activities through mutagenesis and there is scope to extend structural studies to NCS variants other than wild-type NCS. This chapter describes these attempts; the co-crystallisation of intermediate analogues with unnatural substrates or products, studying other enzyme variants and/or proteins

with high similarity but differing activities, and the generation and screening of novel NCS variants.

2.2 Results and discussion

2.2.1 Expression and purification of Δ33*Tf*NCS

Successful crystallographic studies of *Tf*NCS have been previously carried out, in which the 'dopamine-first' mechanism was confirmed by the co-crystallisation of a non-productive reaction intermediate analogue with a construct consisting of predominately the globular fold, ΔN33C196*Tf*NCS (herein referred to as Δ33*Tf*NCS).[30] As well as removing the *N*-terminal signal peptide (1-19), further flexible residues 20-33 and 196-210 have been removed. Therefore, to investigate the activities of the wild-type enzyme for the reactions with non-natural substrates, the enzyme was first expressed and purified based upon the protocol published.[30]

The expression and purification were straightforward, following the reported three-step procedure. Notably, the protein was first isolated by nickel-affinity chromatography (Figure 2.2a), followed by TEV cleavage to remove the His6 tag. To isolate the TEV-cleaved protein and remove TEV protease, another nickel-affinity column was performed, followed by gel filtration to further purify the protein. Complete cleavage of His6-tag was achieved (lanes 2 and 3 of Figure 2.2b). SDS-PAGE analysis of the purification process is shown in Figure 2.2. Pure protein was isolated and concentrated to around 12-13 mg mL^{-1} ready for crystallisation trials. Multiple expressions and purifications of this construct of *Tf*NCS were performed to generate enough protein for multiple rounds of crystallisation trials with various ligands to probe the enzymatic mechanism.

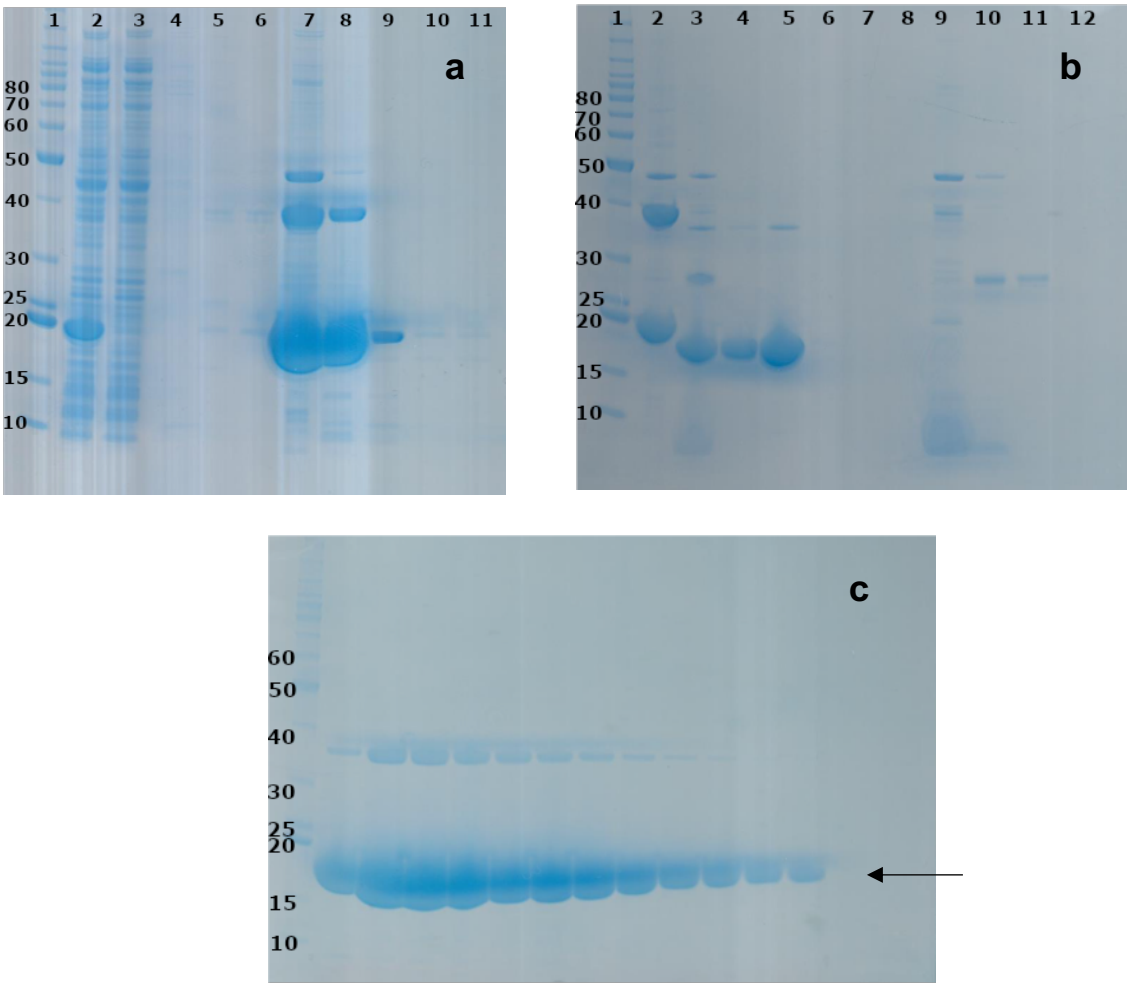


Figure 2.2: **SDS-PAGE analysis of the purification of Δ33*Tf*NCS**. **a.** His-trap purification of Δ33*Tf*NCS at 20.7 kDa. Lanes: 1, Benchmark™ Protein Ladder masses given in kDa. 2, clarified cell lysate. 3, flow through with lysis buffer. 4, wash with 6 CV 20 mM imidazole buffer. 5, wash with 6 CV 40 mM imidazole buffer. 6-11, wash with 500 mM imidazole buffer and collected 4 mL fractions. **b.** His-trap purification of Δ33*Tf*NCS to remove TEV protease and isolated TEV-cleaved Δ33*Tf*NCS at 18.1 kDa. Lanes: 1, Benchmark™ Protein Ladder, masses given in kDa. 2, Pre-TEV cleavage sample (Lane 7 from gel 1, protein at 20.7 kDa) 3, sample loaded onto column. 4, flow through with lysis buffer. 5, 3 CV 20 mM imidazole. 6, 8 CV 20 mM imidazole. 7, 2 CV 40 mM imidazole. 8, 5 CV 40 mM imidazole. 9 and 10, 2 CV 500 mM imidazole. 11, 5 CV 500 mM imidazole. **c.** Gel filtration of Δ33*Tf*NCS: gel of peaks isolated from major peak from Superdex S75 gel filtration. Δ33*Tf*NCS is indicated by an arrow, at 18.1 kDa. Some dimerization of the protein was observed by SDS-PAGE (at 39.2 kDa). Lane 1: Benchmark™ Protein Ladder, masses given in kDa

2.2.2 Improving crystallographic resolution

Initially, structural studies of *Tf*NCS by NMR and X-ray crystallography were performed using a construct lacking the first 29 or 19 *N*-terminal residues, respectively.[27] NCS is predicted to have signal peptides at the *N* and *C*-termini which are non-essential for catalytic activity. Although a high-resolution structure (2.1 Å) of Δ19*Tf*NCS was achieved by Ilari *et al*. when crystallised in the presence of dopamine and 4-hydroxybenzaldehyde (PDB: 2VQ5), the apo structure was gained at 2.7 Å (PDB: 2VNE).[27] Attempts by two previous PhD students, B. Lichman and A. Sula, to gain high resolution structures of the Δ29*Tf*NCS construct with a *C*-terminal His-tag proved challenging, with apo structures gained at resolutions of around 3.5 to 4.0 Å. This suggested that the ten *N*-terminal residues removed were providing an interaction essential for tight crystal packing. The resolution gained was not sufficient to probe the NCS mechanism as the conformations of residue side chains could not be determined.[31] A construct was developed with an *N*-terminal His-tag with a TEV-cleavable linker, so that that His-tag could be removed, however little improvement in resolution was observed. A long *C*-terminal extension was noted during refinement, leading to the development of a novel construct, consisting of solely the globular fold of the protein, named Δ*N*33*C*196*Tf*NCS (also referred to as Δ33*Tf*NCS) removing residues in the terminal signal peptides, 1-33 and 196-210. The construct had an *N*-terminal His-tag with a TEV cleavage site for removal of the His-tag, leading to a more globular fold.

This construct led to high resolution apo structures (*ca*. 2 Å, PDB: 5N8Q) and a co-crystallised structure (1.85 Å, PDB: 5NON) was gained with a non-productive reaction intermediate analogue, to mimic the natural NCS reaction. This structure was used to support the proposed 'dopamine-first' mechanism of action, as discussed further in Section 1.3. Analysis of the structures gained with the Δ33*Tf*NCS construct showed that the *N*-terminal residues could not be resolved (residues 33-39 for PDB: 5NON) or were

not part of the globular fold (38-41 for PDB: 5N8Q). Therefore, to maximise the proportion of globular fold in the construct to improve tight crystal packing and further improve crystallographic resolution, it was hoped that a new construct, removing residues 1 to 41 and 196 to 210 could be used, also with an *N*-terminal His-tag with a TEV-cleavable linker. Unfortunately, no expression of the protein was observed by SDS-PAGE analysis or by western blot using an anti-His-tag antibody. It is possible that the *N*-terminal residues 33-40 are required for stabilisation of the enzyme. Indeed, studies of truncated *Cj*NCS2 constructs by Nishihachijo *et al*. showed that significantly lower levels of expression were observed for Δ42*Cj*NCS2 compared with Δ29*Cj*NCS2.[35]

2.2.3 Initial crystallisation trials of *Tf*NCS

Structural studies of *Tf*NCS in this work were performed using the Δ33*Tf*NCS construct. Initially, studies focussed on reproducing high resolution apo structures of the enzyme to help focus the range of conditions that would be useful for co-crystallisation purposes. Crystallisation trials of the apo protein immediately after purification (at 12.3 mg mL^{-1} in 20 mM Tris, 50 mM NaCl, pH 7.5) were prepared using commercially available 96-well screens, Structure, PACT, JCSG and Index. Although crystallisation was observed in numerous conditions, only those observed in the Structure screen were of appropriate size for data collection. A range of crystals were sent to a synchrotron (Diamond i24); however, diffraction was only observed for three crystals. The crystallisation conditions, resolution and space group of each dataset collected is given in Table 2.1.

Table 2.1: Apo Δ33*Tf*NCS data collection. The crystallisation conditions, dataset resolution and space group are given.

Dataset	*Salt*	*Buffer*	*pH*	*Precipitant*	*Dataset resolution / Å*	*Space group*
1	-	0.1 M sodium citrate	5.6	20% w/v PEG 4000 20% v/v propan-2-ol	4.1	$P3_121$
2	-	0.1 M sodium HEPES	7.5	20% w/v PEG 4000 10% v/v propan-2-ol	2.3	$P3_121$
3	-	0.1 M MES	6.5	12% w/v PEG 20,000	2.0	$P3_121$

A high resolution apo structure was gained from dataset 3, in a novel space group to that previously reported ($P3_121$ vs. $P22_12_1$ for PDB: 5N8Q). To gain mechanistic insights, resolution up to 2.5 Å would be sufficient to resolve the amino acid side chains. A four-corner screen was therefore prepared around the condition of dataset 3 in an attempt to further improve resolution. Numerous crystals were obtained, however there was no improvement.

Previous co-crystallisation and soaking attempts of substrates and products, dopamine (**5**), 4-HPAA (**6**) or (*S*)-norcoclaurine (**7**) by A. Sula and B. Lichman were unsuccessful. No positive difference density was observed in the active site of the enzyme and the resolution was negatively affected.[31] To see whether these results were reproducible and to attempt to understand the acceptance of ketones by *Tf*NCS, co-crystallisation screens were prepared using dopamine (**5**) or one of three spirotetrahydroisoquinoline products (**59-61)**, prepared by Jianxiong Zhao (Figure 2.3) at 10 mM in DMSO with 12.6 mg mL^{-1} Δ33*Tf*NCS.

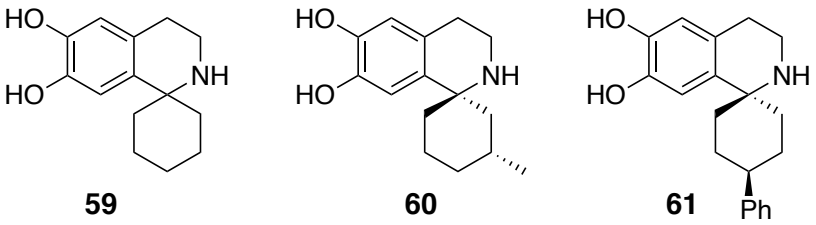


Figure 2.3: Spiro-THIQs used for co-crystallisation. The acceptance of ketones as the carbonyl substrate for a *Tf*NCS-catalysed reaction has led to the production of a range of novel, single isomer spiro-THIQ products.

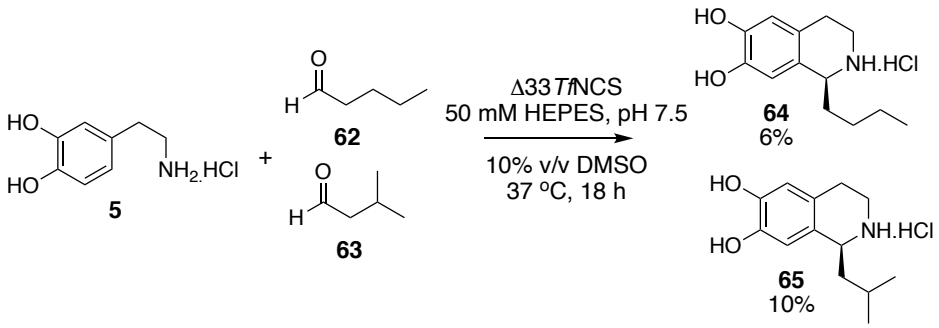
Structure screens were individually prepared for each of the four co-crystallisations (dopamine and **59**-**61**), using 10 mM final concentration of ligand, however few crystals were observed with significant evidence of catecholamine oxidation in the crystallisation wells as evidenced by the precipitation of a brown solid within the crystallisation drop.[137] Three datasets were collected (as given in Table 2.2) but unfortunately, no co-crystallisation was observed. It is likely that dopamine simply diffuses in and out of the enzyme active site and so is not bound sufficiently during crystallisation. The spiro-THIQs are products of an NCS reaction so to drive the NCS reaction, product binding to the active site is probably highly disfavoured. Therefore, in future studies to understand the NCS mechanism beyond the natural substrates, non-productive reaction intermediate analogues were synthesised, mimicking the iminium ion intermediate in the reaction pathway. This approach has already been used successfully to support the 'dopamine-first' mechanism of *Tf*NCS, however had not yet been applied towards non-natural substrates.[30]

Table 2.2: Data collection for Δ33*Tf*NCS co-crystallised with dopamine or spiro-THIQs. The crystallisation conditions, dataset resolution and space group are given.

Ligand	*Salt*	*Buffer*	*pH*	*Precipitant*	*Dataset resolution / Å*	*Space group*
5	-	50 mM KH_2PO_4	-	20% w/v PEG 8000	2.0	$P22_12_1$
59	-	50 mM KH_2PO_4	-	20% w/v PEG 8000	3.2	$P3_121$
60	-	0.1 M MES	6.5	12% w/v PEG 20,000	2.9	$P3_121$

2.2.4 Towards understanding the acceptance of aliphatic aldehydes by *Tf*NCS

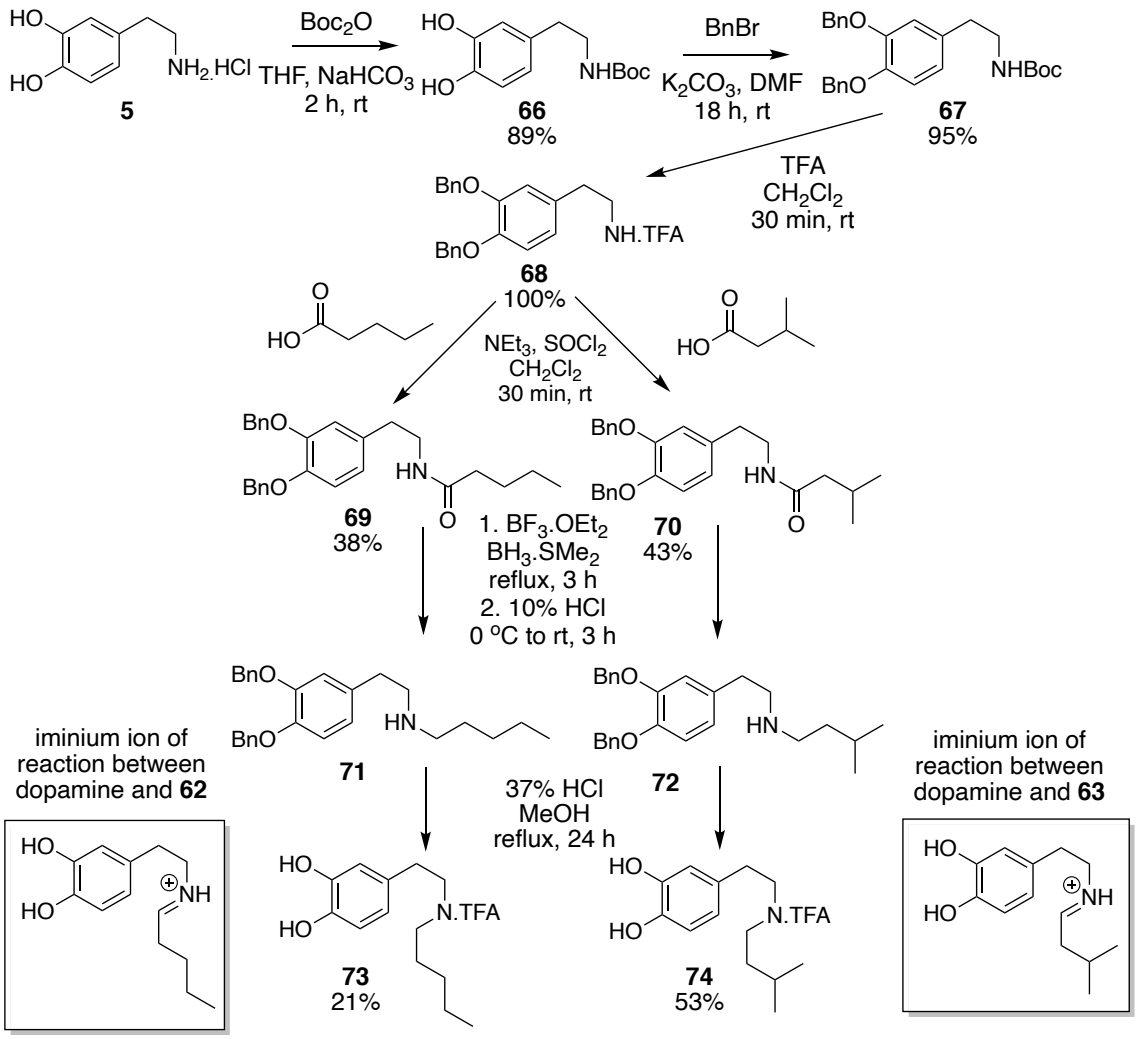
Previous studies of *Tf*NCS by Lichman *et al*. and Ruff *et al*. demonstrated that the aldehyde substrate scope of *Tf*NCS can be extended from benzylic aldehydes towards a range of aliphatic aldehydes, including linear and β-methyl substituted aldehydes.[21,31] Unusually, long alkyl chains on the aldehyde are tolerated, with decanal and citronellal used productively by the enzyme. To understand how *Tf*NCS can accept these aldehydes as substrates, the aim was to use x-ray crystallographic studies with active-site bound non-reactive mimics of reactions between dopamine (**5**) and two aldehydes, one linear (**62**) and one β-methyl substituted (**63**). Enzymatic reactions were performed to confirm the substrate acceptance as neither aldehyde had been tested before (Scheme 2.1). Isolated yields are given, and both were shown to be highly productive as NCS substrates.



Scheme 2.1: *Tf*NCS catalysed reactions performed between dopamine HCl and linear and β-methyl substituted aldehydes. Reactions were performed on a 10 mL scale using 10 mM **5**, 20 mM **62** or **63**, using 0.1 mg mL^{-1} final concentration of Δ33*Tf*NCS. Products (**64** and **65**) were isolated by preparative HPLC (method 1).

Attempts to co-crystallise the THIQ products generated (**64** and **65**) were unsuccessful with only apo datasets collected. Therefore, two secondary amines for co-crystallisation with Δ33*Tf*NCS were designed based upon the iminium ion intermediates of reactions of the Pictet-Spengler reactions towards **64** and **65**, respectively (Scheme 2.2). The synthetic route was similar to that performed by Lichman *et al*.[30] Starting from commercially-available dopamine hydrochloride (**5**), an *O*-benzyl protected analogue (**68**) was generated in high yield (85% in three-steps) via the Boc-protected amine (**66**)

followed by benzyl protection to give **67** and Boc deprotection. The reagents used for the amide coupling step differed from the previous route.[30] Attempts to use DCC (dicyclohexylcarbodiimide) and DMAP (4-dimethylaminopyridine) as a coupling reagent were unsuccessful as the cyclohexylurea by-product co-eluted with the desired amides (**69** and **70**) by thin-layer chromatography. Instead, a one-pot route was used, involving the *in-situ* activation of the carboxylic acid to the acyl chloride.[138] Both amides were obtained in reasonable yield (38% and 43% for **69** and **70**, respectively) after purification by silica column chromatography. The amide reduction steps using borane reagents proceeded to completion in both cases and the secondary amines (**71** and **72**) were taken through to the benzyl deprotection step without further purification. Benzyl deprotection was achieved through refluxing under acidic conditions and the desired mimics for co-crystallisation were isolated by preparative-HPLC as the TFA salts in 21% and 53% yield for **73** and **74**, respectively. The synthetic route is summarised in Scheme 2.2. Both mimics (**73** and **74**) were prepared at 200 mM in DMSO ready for co-crystallisation.



Scheme 2.2: Synthetic route towards two aliphatic aldehyde reaction mimics for co-crystallisation with Δ33*Tf*NCS. The final products are isolated as the TFA salt after purification by preparative HPLC.

Many crystals were obtained from commercial screens and optimisation was performed for numerous conditions. However, convincing positive difference density was never observed in the enzyme active site. Poor resolution was observed (3 – 4 Å) compared with apo structures gained under the same conditions (*ca.* 2 Å) suggesting that the presence of the ligands disrupted tight packing in the crystal lattice. Both mimics are also significantly more conformationally flexible than the one used previously (**16**). The lack of a second aromatic ring also means that there is less electron density. Very high occupancy, and a single ligand orientation would be required for significant difference density and accurate ligand placement. It is known that the NCS mechanism is highly dynamic by NMR spectroscopic studies and that iminium ion reorientation occurs during

the NCS mechanism, so co-crystallising the ligand in a single orientation proved challenging, hence the large number of datasets collected. Soaking of the apo crystals was also unsuccessful, with crystals destroyed upon addition of the mimics. Therefore, for future crystallographic investigations, mimics were prepared with two aromatic rings present, as discussed in Sections 3.2.6, 5.2.13 and 6.2.4.

2.2.5 Sacred lotus NCSs

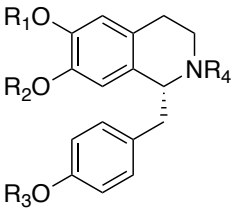
To date, two variants of NCS have been most widely explored, *Tf*NCS (isolated from *Thalictrum flavum*) and *Cj*NCS2 (isolated from *Coptis japonica*). A range of phenethylamines and aldehydes beyond the natural substrates have been shown to be accepted by both enzymes to generate a range of novel (1*S*)-THIQs in a single-step. Single enantiomer syntheses are valuable as often only a single enantiomer possesses the desired biological activity, with the other enantiomer enhancing unfavourable side-effects, possessing other biological activities or simply being inactive.[139] Therefore, being able to access all isomers of a small molecule drug candidate is highly valuable.

NCSs have proved to be highly effective for generating (1*S*)-THIQs in facile, single-step syntheses, therefore a variant or mutant capable of generating the corresponding (*R)*-enantiomers would be highly valuable. In all cases, norcoclaurine in plants has been identified as the (*S*)-enantiomer, with a single exception from the sacred lotus, *Nelumbo nucifera.*[140] There is scientific interest in the identification and isolation of active components of *N. nucifera* as extracts of various parts of the plant have been used in traditional medicine to treat a variety of disorders, from smooth muscle and uterine relaxation[141] to dyspepsia.[47]

A range of interesting, novel benzylisoquinoline alkaloids (BIAs) have been isolated from *N. nucifera* and can be divided into three different subclasses: aporphine, 1-benzylisoquinoline, and bisbenzylisoquinoline (from the oxidative coupling of two 1-

BIAs). A summary of some of the known (1*R*)-aporphine and (1*R*)-benzylisoquinoline alkaloids isolated are given in Figure 2.4.

(1*R*)-benzylisoquinoline	R_1	R_2	R_3	R_4
Norcoclaurine* (**75**)	H	H	H	H
Coclaurine	Me	H	H	H
N-methylcoclaurine	Me	H	H	Me
Norarmepavine	Me	Me	H	H
Armepavine*	Me	Me	H	Me



(1*R*)-aporphine	R_1	R_2	R_3
Caverine	Me	H	H
Asimilobine	H	Me	H
Lirinidine	Me	H	Me
O-Nornuciferine	H	Me	Me
N-Nornuciferine	Me	Me	H
Anonaine	$-CH_2-$		H
Roemerine	$-CH_2-$		Me

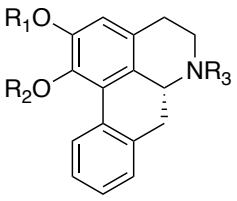


Figure 2.4: Various (1*R*)-BIAs isolated from *Nelumbo nucifera*. * indicates that the *S* enantiomer has also been isolated. Other THIQ alkaloids have also been isolated from *N.* nucifera but the absolute stereochemistries have not been determined.

The aporphine alkaloids isolated from *N. nucifera* are interesting due to the lack of substitution on the benzylic moiety and *R* stereochemistry at the C-1 position. The common intermediate to all BIAs is widely regarded to be (*S*)-norcoclaurine (**7**), with alkaloids with the opposing stereochemistry at C-1 often derived from (*R*)-reticuline (**21**) so the biosyntheses of these alkaloids is clearly unique. **21** and 1-benzylisoquinoline

alkaloids lacking *para*-hydroxy or -methoxy substitution on the benzylic ring have also never been isolated from plants although this does not rule them out as intermediates that are later modified. Biosynthetic possibilities include the involvement of a dehydroxylase enzyme (although rarely reported)[142] to remove the benzylic *para*-hydroxyl group with a route via *R*-norcoclaurine (**75**), the presence of a Pictet-Spenglerase that catalyses the reaction between dopamine (or an analogue) and phenylacetaldehyde or an epimerase or other stereoselective enzyme involved to generate the THIQ moiety. The aporphine scaffold is likely then generated by an oxidative coupling between the two aromatic rings, likely catalysed by a P450 enzyme.[143]

A range of 1-benzylisoquinoline alkaloids have also been identified in *N. nucifera*, with structures based upon norcoclaurine, with various methylation patterns on the heteroatoms (Figure 2.4). A unique feature of these alkaloids is that again, (*R*)-stereochemistry at C-1 was observed in some cases. Different enantiomers also appear to localise in different parts of the plant. The presence of solely the (*R*)-enantiomer of norcoclaurine in the seed embryo was first identified by Koshiyama *et al*. in 1970, by comparison of the optical rotation found of the compound isolated to those of known stereochemistry. This was also confirmed by HPLC analysis based on derivatisation with a chiral, fluorescent agent.[144]

The biosynthetic origin of these alkaloids is again unknown, although the presence of seven isoforms of NCS (*Nn*NCS1-7) in the *N. nucifera* genome means that it is highly likely that one of these enzymes is involved. Multiple sequence alignment of the NnNCSs, *Cj*NCS2 and *Tf*NCS is shown in Figure 9.56. However, the observation of both enantiomers of norcoclaurine, localised in different parts of the lotus, means that elucidating the biosynthetic origins is challenging. It is possible that a racemic reaction occurs followed by *in planta* transport systems to localise the varying enantiomers, an epimerase enzyme is also present to convert (*S*)-norcoclaurine (**7**) to the corresponding

R enantiomer or that one of the NCS variants present is capable of biosynthesising solely (*R*)-norcoclaurine (**75**).

Seven isoforms of *Nn*NCS have been identified in the sacred lotus genome and categorised as being one of two subfamilies of NCS, NCS*I* and NCS*II*.[46] NCS*I* genes are found in BIA-expressing plants whereas NCS*II* genes are found in plants which do not express BIAs. Both *Nn*NCS2 and *Nn*NCS6 were deemed to be pseudogenes due to a frameshift from a single nucleotide deletion and partial coding sequences, respectively. Phylogenetic analysis also deemed that *Nn*NCS4 was part of the non-BIA producing NCS*II* clade. Interestingly, unlike other encoded NCSs, all *Nn*NCSs did not possess terminal signal peptides thus suggesting an alternative subcellular localisation, the cytoplasm rather than the vacuole.[18]

Therefore, it was possible that one of the four remaining *Nn*NCSs may be capable of catalysing (*R*)-norcoclaurine (**75**) so were considered for investigation. Unfortunately, there is no detail on enzyme expression in the seed embryo, where **75** is localised. Most significant alkaloid production in the leaves was correlated with *Nn*NCS7 expression. The sequence similarities of the four variants with *Tf*NCS and comparisons of the key active site residues are given in Tables 2.3 and 2.4, respectively.

Table 2.3: Sequence similarity and identity between *Tf*NCS and the four *Nn*NCSs.

***Nn*NCS variant**	**Identity (%)**	**Similarity (%)**
***Nn*NCS1**	30	46
***Nn*NCS3**	30	45
***Nn*NCS5**	32	47
***Nn*NCS7**	29	45

Table 2.4: Variations in key active site residues between *Tf*NCS and the four *Nn*NCSs.

Active site residue in *Tf*NCS	K122	E110	Y108	D141	M97	A79	F99	L72	L68	F80	F112
***Nn*NCS1**	K	E	W	V	I	V	L	F	L	Y	F
***Nn*NCS3**	K	E	W	I	I	V	L	F	L	Y	F
***Nn*NCS5**	K	E	W	D	I	V	L	V	L	Y	F
***Nn*NCS7**	K	E	Y	I	V	V	L	C	L	I	F

All *Nn*NCSs had high sequence similarity with *Tf*NCS (Table 2.3 and Figure 9.56). Active site residues were compared with *Tf*NCS (Table 2.4) to help determine whether catalytic activity would be retained. Key active site residues, known to be essential for catalytic activity, K122 and E110 were retained (Table 2.4). K122 is responsible for catechol binding and E110 is thought to be important for iminium ion formation and orientation.[28] Interestingly, D141 was not conserved across all variants. This residue is implicated in iminium ion formation and activity is significantly reduced in *Tf*NCS when mutated to D141E or D141N, implying the importance of residue chain length and charge for successful catalysis.[28] The putative gatekeeper residue, F112 is also conserved. Y108 is only conserved in *Nn*NCS7 and although it is thought to aid aldehyde binding to the active site, the phenol is not essential for catalytic activity; Y108F is also a highly productive catalyst.[31]

Interestingly, residue L72 was not conserved. Computational studies of *Tf*NCS by Sheng *et al*.[33] demonstrated that the pathway towards both enantiomers of norcoclaurine are favoured until the final deprotonation step of the final quinone intermediate to give the THIQ product. Moreover, the route towards (*R*)-norcoclaurine is disfavoured in part due to an unfavourable steric clash between the *para*-hydroxyl of the benzylic moiety and L72. The alteration, L72F increases steric bulk at this position,

thus suggesting that the steric repulsion would be enhanced, and *R*-product would be disfavoured. However, the variation L72V in *Nn*NCS5 does decrease steric bulk but it is unknown if this creates enough space in the active site to avoid the steric repulsion and favour *R*-product formation. To avoid formation of the *S*-enantiomer as well, a steric clash on the other side of the active site would also be required. The residue change to L72C (in *Nn*NCS7) decreases steric bulk but also side chain flexibility.

Therefore, two NCSs isolated from the sacred lotus, *Nn*NCS5 and *Nn*NCS7 were investigated due to apparent decreased steric bulk at position 72 and correlations with increased alkaloid production, respectively. Due to the lack of signal peptides, no truncations of the termini were performed when designing synthetic genes for protein expression. Synthetic genes were designed and subcloned into pET29a(+), then used to transform BL21(DE3). This work was performed by Prof. John Ward. A construct was developed with an *N*-terminal His-tag with a TEV cleavage site because if a productive enzyme was found, crystallographic studies would be of interest and this construct was promising with *Tf*NCS.[30] Expression was performed using the same protocol as for the expression of *Tf*NCS (wild-type and variants) as it was hoped that due the high sequence homology observed between the NCSs, this would be successful. Multiple sequence alignment of the *Nn*NCSs and *Tf*NCS is shown in Figure 9.56. SDS-PAGE analysis of the whole-cell and lysed supernatant of the whole-cell expression cultures (Figure 2.5) lacked the presence of an over-expressed protein. Therefore, western blot analysis was performed, using a fluorescent probe to detect for the His-tag.

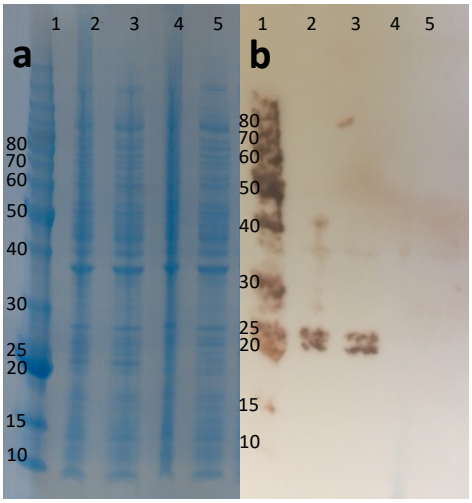


Figure 2.5: SDS-PAGE and western blot analysis of *Nn*NCS5 and *Nn*NCS7 expression. Expected protein sizes are: *Nn*NCS5 at 21.7 kDa and *Nn*NCS7 at 21.1 kDa. Expression was performed using the same method as reported for the generation of *Tf*NCS (1 mM final concentration IPTG to induce followed by incubation at 25 ºC). **a.** SDS-PAGE analysis. **b.** Western blot analysis using an Anti His-tag antibody. Lanes: 1, Benchmark™ Protein Ladder (masses given in kDa). 2, *Nn*NCS5 whole cell. 3, *Nn*NCS5 supernatant. 4, *Nn*NCS7 whole cell. 5, *Nn*NCS7 supernatant.

No expression was observed for *Nn*NCS7. Two bands were observed for *Nn*NCS5 expression, perhaps indicating the presence of a degradation product and very low levels of protein expression were observed. The pure protein was isolated to minimise the observation of the non-enzymatic background reaction. Nickel-affinity chromatography was performed followed by TEV-cleavage, nickel-affinity chromatography to remove TEV protease and size-exclusion chromatography. SDS-PAGE analysis of the fractions isolated from each purification step is given in Figure 2.6.

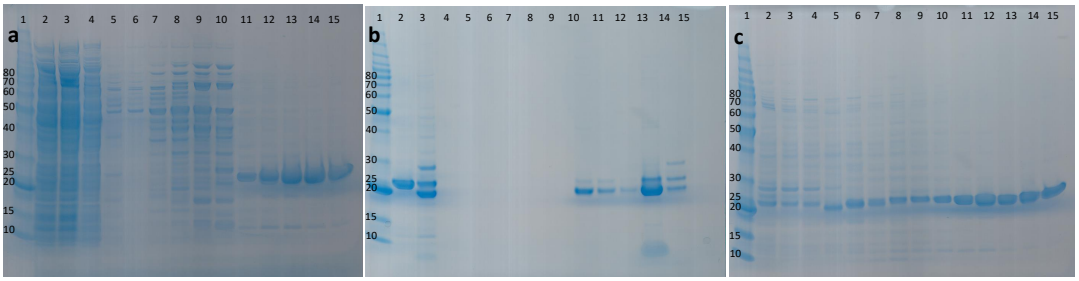


Figure 2.6: SDS-PAGE analysis of *Nn*NCS5 purification. CV = column volume. **a**. SDS-PAGE analysis of nickel-affinity chromatography purification of *Nn*NCS5. Lanes: 1, Benchmark™ Protein Ladder (masses given in kDa). 2, clarified cell lysate loaded onto column. 3, flow through with lysis buffer (20 mM imidazole). 3 – 6, 20 mM imidazole buffer wash, 2 CV fractions. 7 – 9, 40 mM imidazole buffer wash, 3 CV fractions. 10 – 15, 500 mM imidazole buffer washes. 1 CV fractions. **b.** SDS-PAGE analysis of second nickel-affinity chromatography purification of *Nn*NCS5 after His-tag removal to remove TEV protease. Lanes: 1, Benchmark™ Protein Ladder (masses given in kDa). 2, *Nn*NCS5 sample before TEV cleavage at 21.7 kDa. 3, *Nn*NCS5 sample loaded onto His-trap after TEV cleavage, protein expected at 19.1 kDa. 4, flow through with 20 mM imidazole buffer. 5 - 7, 3 CV 20 mM imidazole buffer wash. 8 - 12, 3 CV 40 mM imidazole buffer wash, 13 – 15, 3 CV 500 mM imidazole buffer wash. **c.** SDS-PAGE analysis of size exclusion chromatography purification of *Nn*NCS5, showing isolated protein at 19.1 kDa. Lanes: 1, Benchmark™ Protein Ladder (masses given in kDa). 2 – 15, 3 mL fractions collected across the major protein containing fractions eluted from the column.

The first nickel-affinity purification of *Nn*NCS5 removed most of the other protein impurities and fractions were pooled corresponding to lanes 11-15 by SDS-PAGE analysis (Figure 2.6a). TEV cleavage of the His-tag did not go to completion as shown in Figure 2.6b, and since no issues had been observed before during His-tag cleavage of *Tf*NCS, it was hypothesised that in *Nn*NCS5, the His-tag is less accessible, hindering access by the TEV protease. Unusually, despite removal of the His-tag from some of the protein sample, *Nn*NCS did not elute in the flow through or 20 mM imidazole buffer washes and instead eluted later in 40 mM and 500 mM imidazole washes. A reasonably pure sample of protein was achieved after size exclusion chromatography (Figure 2.6c) with fractions corresponding to lanes 7 – 15 pooled. A low yield of protein was isolated, 1 mg from 3 L of culture volume. This is significantly lower than yields of *Tf*NCS where 100 – 200 mg of pure protein was regularly isolated from a single litre of culture. Poor expression could be due to toxicity or instability of the *Nn*NCS5. There are also clearly

issues associated with having an *N*-terminal His-tag, so in future a *C*-terminal tag could be considered instead. The reported *C*-terminal of *Nn*NCS5 contains five consecutive serine residues, which could be due to sequencing errors. Therefore, truncated constructs of *Nn*NCS5 could also be considered.

Purified *Nn*NCS5 was concentrated to 1 mg mL^{-1} in buffer (20 mM Tris, 50 mM NaCl, pH 7.5) and stored at -80 °C or used directly for enzymatic reactions. Assays were then performed to determine whether the isolated enzyme was active. Reactions were considered based upon reactions with *Tf*NCS because, due to the high sequence homology of the two NCSs, if active, it was hoped that *Nn*NCS5 would have similar reactivity to *Tf*NCS. The two substrates chosen for assays were dopamine (**5**), the natural NCS phenethylamine substrate and phenylacetaldehyde (**76**), a commercially available and less oxidatively sensitive analogue of the natural substrate, 4-HPAA (**6**). Product formation was monitored by RP-HPLC analysis and confirmed by LC-MS. Background, control reactions were performed by substituting the protein sample for buffer (20 mM Tris, 50 mM NaCl, pH 7.5).

Unfortunately, using 1 mM or 10 mM of either substrate, with or without 10% v/v DMSO as co-solvent and at 37 °C or 4 °C, no product formation was observed above that observed for the non-enzymatic background reaction. A 0.4 mg mL^{-1} final concentration of enzyme was used which should be sufficient to demonstrate if the enzyme was active. Therefore, due to a combination of very low expression yields and no (or potentially trace) activity, further optimisation of *Nn*NCS5 expression and assay conditions was not performed.

*Nn*NCS7 expression was attempted following the same protocol as *Nn*NCS5 and *Tf*NCS expression however, no expression of *Nn*NCS7 was observed. Therefore, expression trials were performed, varying IPTG concentrations (0.1 mM or 0.5 mM final)

used for induction and two different incubation temperatures post-induction (18 °C and 26 °C) were attempted. Protein expression could not be determined by SDS-PAGE analysis; therefore, the presence of *Nn*NCS7 was identified by western blot analysis using an Anti His-Tag Antibody to detect protein containing a His-tag (Figure 2.7). To determine whether the expressed protein was soluble, samples were prepared after lysis (whole cell), the isolated supernatant after centrifugation (supernatant) or the resuspended cell pellet (pellet) and were analysed.

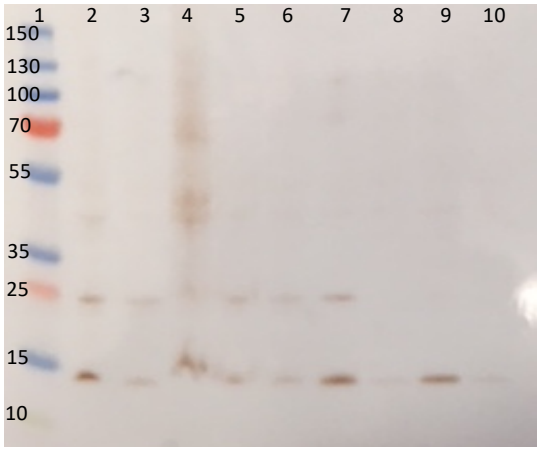


Figure 2.7: Anti His-Tag western blot analysis of *Nn*NCS7 expression trials. Protein expected at 21.1 kDa. Lanes: 1, Thermo Scientific™ PageRuler™ Plus Prestained Protein Ladder. 2, 0.5 mM IPTG, 18 °C, whole cell. 3, 0.5 mM IPTG, 18 °C, supernatant. 4, 0.5 mM IPTG, 18 °C, pellet. 5, 1 mM IPTG, 18 °C, whole cell. 6, 1 mM IPTG, 18 °C, supernatant. 7, 0.5 mM IPTG, 26 °C, whole cell. 8, 0.5 mM IPTG, 26 °C, supernatant. 9, 1 mM IPTG, 26 °C, whole cell. 10, 1 mM IPTG, 26 °C, supernatant.

Presence of the desired protein (at 21.2 kDa) was observed for all expression conditions, however in all cases, a degradation product was observed at around 15 kDa, thus indicating the instability of the protein. Induction with 1 mM IPTG and incubation at 26 °C resulted in only degraded protein. Therefore, to isolate *Nn*NCS7 to determine whether it was active, the enzyme was expressed with 0.5 mM IPTG and incubation at 18 °C post-induction. To purify the protein, nickel-affinity chromatography was performed, with SDS-PAGE analysis of fractions eluted shown in Figure 2.8. There was some evidence of oligomerisation and the protein sample obtained was impure.

Due to the likely instability of the protein, further purification steps (TEV cleavage, secondary His-trap and size-exclusion chromatography) were not performed. A poor yield of impure protein (0.2 mg from 3 L culture volume) was obtained

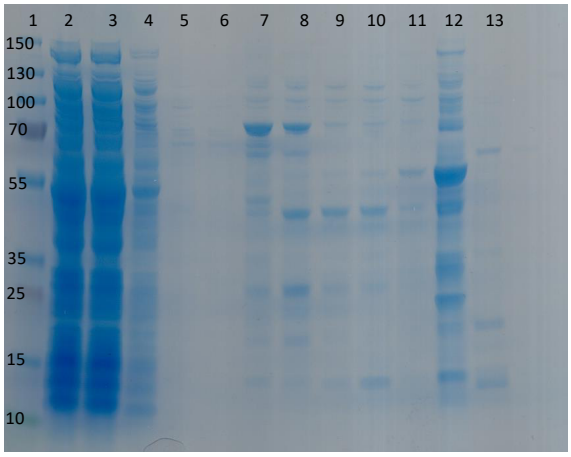


Figure 2.8: SDS-PAGE analysis of *Nn*NCS7 purification by nickel affinity chromatography. Protein expected at 21.1 kDa. CV = column volume. Lanes: 1, Thermo Scientific™ PageRuler™ Plus Prestained Protein Ladder. 2, lysate loaded onto His-trap. 3, flow-through from column. 4 – 6, 5 CV of 20 mM imidazole buffer. 7 – 8, 2 CV 40 mM imidazole buffer. 9, 5 CV 40 mM imidazole buffer. 10, 2 CV 80 mM imidazole buffer. 11, 5 CV 80 mM imidazole buffer. 12, 2 CV 500 mM imidazole buffer. 13, 5 CV 500 mM imidazole buffer.

The isolated protein (sample in lane 12 of Figure 2.7) was concentrated and prepared in buffer as for *Nn*NCS5 and enzymatic reactions were performed using a range of conditions, analogous to those used to determine the biocatalytic potential of *Nn*NC5. Unfortunately, *Nn*NCS7 also showed no reaction above the background reaction using all conditions tested. It is unknown why such poor stability and expression is observed with these two sacred lotus NCSs, considering the high homology to other NCS variants. It is possible that the introduction of long *N*-terminal extensions, from the incorporation of a His-tag and a TEV cleavage site are poorly tolerated and disrupt the proteins.

Aligning the sequences of the two *Nn*NCSs investigated and *Tf*NCS (Figure 9.56) also revealed an unusual feature of the *Nn*NCS5 sequence; five sequential serine residues at the *C*-terminus and the lack of seven residues at the *N*-terminus. In *Nn*NCS7, the

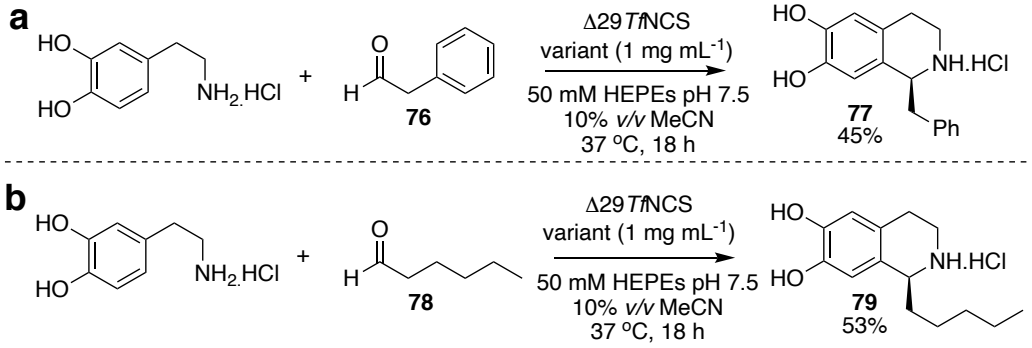
sequence is truncated at the *C*-terminus, however studies of *Tf*NCS have shown that the C-terminal unstructured region is not essential for catalytic activity. In future attempts to isolate these enzymes, truncations of C-terminal residues and C-terminal His-tagging could be explored. There is also a possibility that the five isoforms of NCS isolated from *N. nucifera* have evolved to become inactive, with both enantiomers of norcoclaurine generated via a phosphate-mediated Pictet-Spengler reaction[15] and each enantiomer selectively transported to and accumulating in different parts of the plant.[140]

2.2.6 Investigations of double point variants of *Tf*NCS

A range of double point variants of *Tf*NCS have been generated by a previous PhD student, Benjamin Lichman with the aim of altering the regioselectivity of the Pictet-Spengler reaction towards *ortho*-THIQs rather than the *para*-THIQs generated by the wild-type enzyme. Eight different variants were generated by site-directed mutagenesis, moving key active site residues with the aim to alter the binding location of the *meta*-hydroxyl of the phenethylamine substrate used. The aim was to favour formation of the spirocyclic intermediate via activation of the *ortho*-hydroxyl group. Mutations of K122 (to which the catechol binds) towards non-reactive residues was performed and changing other active site residues might mean that the amine substrate would be bound in a different orientation in the active site and perhaps alter catalytic activity. Some screening had been previously performed by B. Lichman, investigating a range of ketones and an α-keto acid as the carbonyl substrate, however no or trace conversions were observed.[31]

Investigation of these eight double point variants of Δ29*Tf*NCS were investigated as part of a collaborative project with another PhD student, Yu Wang, whereby the aim of the project was to further investigate NCS activity and alter reactivity through rational design of the enzyme. SDS-PAGE analysis of the double point variants is given in the PhD thesis of Yu Wang.[145] The enzymes were expressed using the standard protocol

for *Tf*NCS expression and all were found to be highly expressed. Enzymatic reactions were performed with desalted cell lysate to remove any phosphate or other salts which may be catalysing a racemic, non-enzymatic Pictet-Spengler reaction.[15] Due to the movement of key residue, K122 in the active site, and thus the potential for significant lower enzymatic activity, higher concentrations of enzyme were used in reactions compared with conditions usually used (1 mg mL^{-1} vs. 0.1-0.2 mg mL^{-1}). Screening was performed of the reactions between dopamine and phenylacetaldehyde (**76**) to give **77**, or hexanal (**78**) to give **79**, respectively (Scheme 2.3). These aldehydes were chosen as analogues of the natural substrate, 4-HPAA (**6**) and to determine the activities with linear aldehydes. Peak areas by analytical RP-HPLC (method 1, Figure 2.8) above the background reaction (substituting enzyme sample for buffer instead) indicated a productive biocatalytic reaction and chiral HPLC analysis (method 5, Figure 2.8) using reported conditions was used to determine the stereoselectivity of the reactions.[117] Conversions of reactions with the eight, double-point variants with either **76** or **78** (Scheme 2.3) as the aldehyde substrate are given in Figure 2.9.



Scheme 2.3: Reactions performed to assess the activity of the double point variants of Δ29*Tf*NCS. a. Reaction between dopamine and **76**. **b.** Reaction between dopamine and **78**. Yields given are for isolated yields after a 10 mL scale reaction using 10 mM **5** and 20 mM **76** or **78**, with wild-type Δ29*Tf*NCS (0.1 mg mL^{-1}).

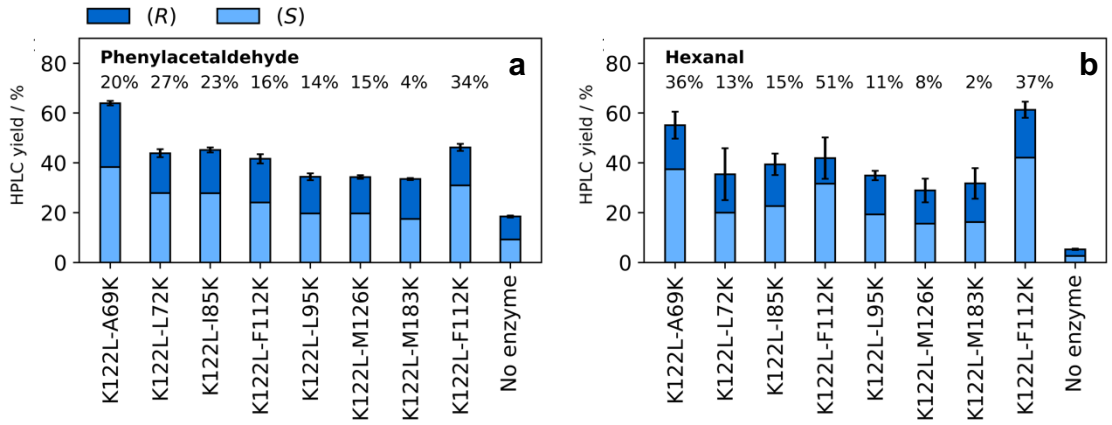


Figure 2.9: Conversions and enantiomeric excesses of reactions involving double-point variants of *Tf*NCS. Reactions were performed using standard conditions between dopamine (10 mM) and **a**. phenylacetaldehyde (**76**) or **b**. hexanal (**78**) (20 mM), catalysed by a range of double point variants of Δ29*Tf*NCS (at 1 mg mL^{-1} using desalted cell lysate). Conversions were determined based upon product formation compared with calibration curves of isolated product. Enantiomeric excesses for each reaction condition are given above each bar as percentages and were determined by chiral HPLC (method 5).

Compared with reactions with involving wild-type *Tf*NCS, conversions and enantiomeric excesses were significantly poorer. With both aldehyde substrates, using 0.1 mg mL^{-1} wild-type *Tf*NCS, near complete conversions were observed in high enantiomeric excess (>95%).[31] For reactions between dopamine (**5**) and **76** (Figure 2.9a), conversions above the background were observed with all double-point variants and in all cases an excess of the (*S*)-enantiomer was observed by chiral HPLC, albeit with poor e.e.s. Poor selectivity was likely due to a combination of a high background reaction and perhaps some racemic reaction associated with the protein sample. This was also observed with reactions involving the benzaldehydes (Section 5.2) although it is possible here that by altering these residues, that instead of providing catalytic activity, the active site instead acts as a hydrophobic pocket. The shape of the active site therefore has some influence over the selectivity of the reaction and enhances turnover compared with the background reaction. The activity of the enzyme was however greatly decreased with significantly lower conversions observed (<64% with

all double point variants with 1 mg mL^{-1} *Tf*NCS) compared with the wild-type enzyme (>95% with 0.1 mg mL^{-1}).

Similar results were also observed for reactions with **78** and in all cases, significant reactivity above the background was observed, along with stereoselectivity in all cases (Figure 2.9b). Overall, the best selectivities were observed when using K122L-F122K as the catalyst (e.e.s of 34% and 37% for **76** and **78**, respectively). F112 neighbours K122 in the active site and is thought to act as a 'gatekeeper residue', influencing substrate acceptance.[30] Perhaps as the two residues are nearby, this means that the catechol can still bind in the active site in a similar location so reactivity is hardly altered. Clearly the presence of an aromatic residue buried in the active site is important although it is unclear why the reactivity varies greatly between using the benzylic and linear aldehyde as substrates. Perhaps enhanced reactivity was observed for reactions with hexanal (**78**) rather phenylacetaldehyde (**76**) as **78** is more conformationally flexible and less bulky, so there is more space in the active site for substrate maneuver. It is particularly interesting that reactivity was observed, with retained stereo- and regioselectivity for all these mutations, considering that the key lysine residue (K122) is moved far from the original location in some cases, even into the entrance/exit of the active site (K122L-M126K). The activity of the enzyme was however greatly decreased with poor conversions (<64% with 1 mg mL^{-1} *Tf*NCS) observed compared with the wild-type enzyme (>95% with 0.1 mg mL^{-1}).

2.2.7 Rational design of NCS towards altering reaction stereoselectivity

Density functional theory calculations of *Tf*NCS, to investigate the reaction mechanism and stereoselectivity indicated the role of L72 in producing solely the (*S*)-product. As previously discussed in Section 1.3.4, it was hypothesised that the reaction pathways towards either enantiomer are favoured until the final quinone intermediate, where an unfavourable steric repulsion is observed between the benzylic moiety of the (*R*)-

quinone and the side chain of L72, therefore meaning that deprotonation of the (*R*)-quinone by E110 is disfavoured. Since the all reactions are reversible except for this final deprotonation step, this means that the formation of (*S*)-norcoclaurine is favoured. Therefore, it was hoped that a rational design of *Tf*NCS could be performed, minimising the unfavourable interaction of the (*R*)-quinone, by reducing steric bulk at position 72, towards residues with smaller side chains, valine and alanine. This may help generate racemic or (*R*)-selective *para*-THIQ products. To disfavour deprotonation of the *S*-quinone, it was hypothesised that a steric clash between the benzyl moiety and an active site residue would be required. Residue, Y108 appeared to be on the other side of the active side compared to L72, and the quinone intermediate so providing extra steric bulk at this position could result in a steric clash and thus disfavour *S*-quinone formation.

Based on this hypothesis four single point variants of Δ29*Tf*NCS were generated; L72A, L72V, Y108S and Y108W. Cloning and expression of these variants was performed by Yu Wang. Initially, only the single point mutations were considered so that any changes in activity could be directly correlated to each mutation. Enzymatic assays were performed under standard conditions for reactions between dopamine and phenylacetaldehyde or hexanal (like Scheme 2.3). Reactions were performed at higher enzyme concentrations (1 mg mL^{-1}) as before, as it was expected that enzyme activity would be negatively affected. Enzyme samples were prepared using desalted cell-lysate to remove any phosphates which could contribute to a background, racemic Pictet-Spengler reaction. Conversions and enantiopurities of the THIQ products generated are given in Figure 2.10.

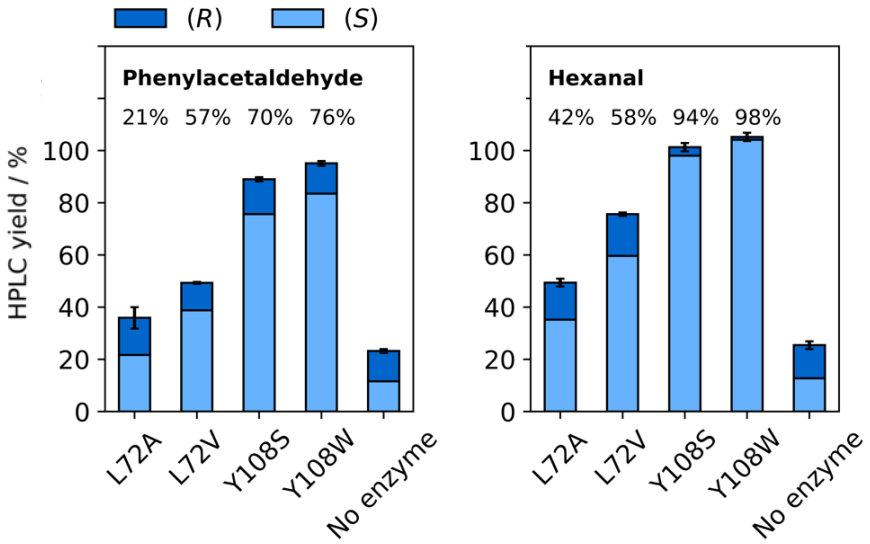


Figure 2.10: Conversions of reactions using single point variants of Δ33*Tf*NCS at positions 72 and 108. Reactions were performed using 1 mg mL^{-1} enzyme under standard reaction conditions between dopamine (**5**) and phenylacetaldehyde (**76**) or hexanal (**78**). Conversions were determined by product formation calibration curves and enantiomeric excesses determined by chiral HPLC (method 5) are given above each bar and enantiomeric ratios are shown by split bars.

For reactions with all four variants and both substrates, selectivity towards the (*S*)-enantiomer of product was observed (Figure 2.9). The amount of (*R*)-product observed was comparable to that observed in the background reaction, unless conversions were at or approached completion (reactions involving hexanal and Y108S/W, Figure 2.9). In all cases, higher e.e.s were observed using hexanal as the aldehyde substrate. Activity was poor with both substrates using the L72 variants and the stereoselectivity of the reaction was unaltered. Conversions decreased as the side chain was shortened, from valine to alanine. Perhaps this is because this residue is important to provide hydrophobic interactions with the substrates rather than being involved in a steric repulsion with the aldehyde side chain that affects the stereochemical outcome of the reaction.

Y108W had high activity, with complete conversions observed with both substrates. To compare the activity of Y108W to the wild-type enzyme, enzyme reactions were performed at a range of different enzyme concentrations between dopamine (**5**) and

hexanal (**78**). **78** was chosen as the aldehyde substrate because for reactions with 1 mg mL^{-1} enzyme (Figure 2.11), the Y108W was highly productive; complete conversion to the product, **79** was observed in 98% e.e. It was shown that the activity of Y108W was approximately half of that of the wild-type enzyme. This resulted in poorer enantiopurities of the THIQ (**79**) generated, likely due to an increased proportion of background-generated product present.

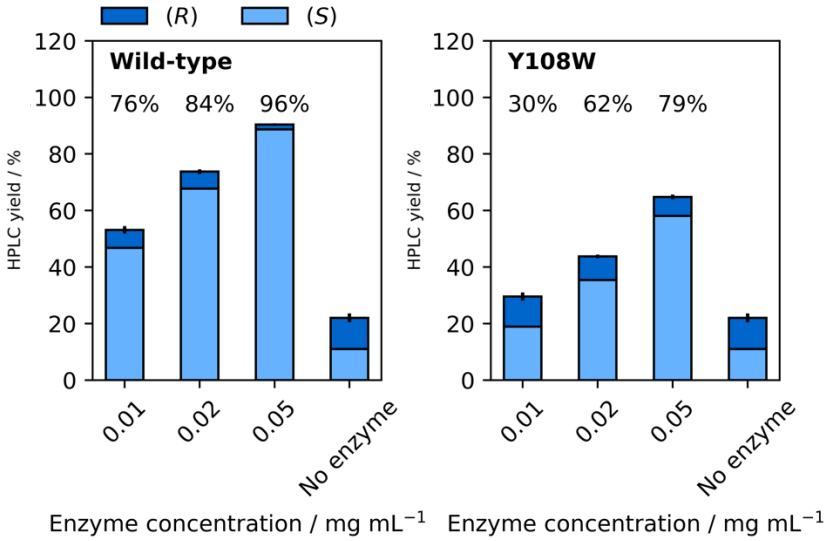


Figure 2.11: Conversions of NCS-catalysed reactions between dopamine and hexanal using wild-type or Y108W Δ33*Tf*NCS at varying concentrations. Enantiomeric excesses of the products generated are given above each bar and shown as an enantiomeric ratio by split bar. Enantiopurities were determined by chiral HPLC method 5.

The mutation of tyrosine to tryptophan also increases the hydrophobicity of the pocket. Smaller chain aldehydes (such as acetaldehyde (**45**)) are poorly accepted as substrates by *Tf*NCS. Therefore, enzyme catalysed reactions between dopamine and acetaldehyde, catalysed by Y108W, were investigated. However, no conversions were observed by RP-HPLC nor LCMS.

Double point variants of *Tf*NCS based upon L72A/S and Y108S/W were therefore not generated because L72 did not appear to have a role in determining reaction selectivity and mutations at this position instead greatly reduced enzyme activity. Instead, another approach was considered, taking advantage of the fact that *Tf*NCS has two entrances

to the active site (Figure 2.12). During the NCS mechanism, the aldehyde side chain sits in a cleft near residues L76, A79, F80 and M97. On the other side of the active site, there is another cleft, flanked by the side chains of Y131 and P179, and the loop 173 – 176. This pocket is inaccessible to the aldehyde side chain, as it is blocked by the side chains of Y108 and P179.

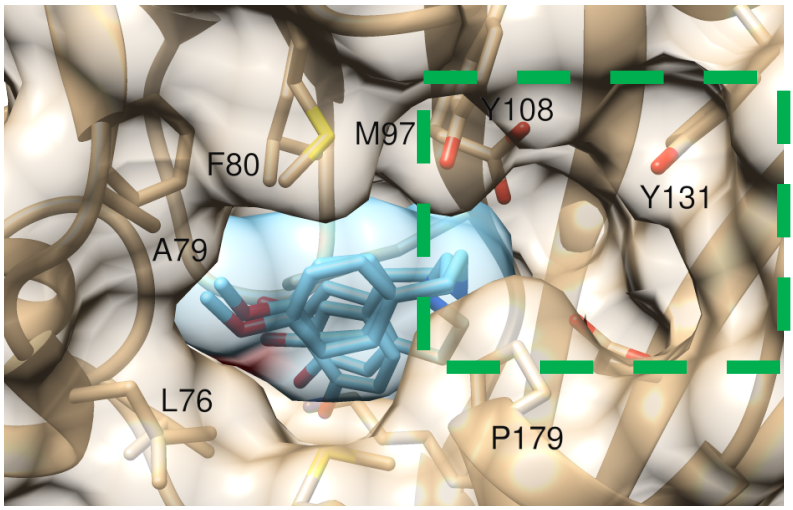
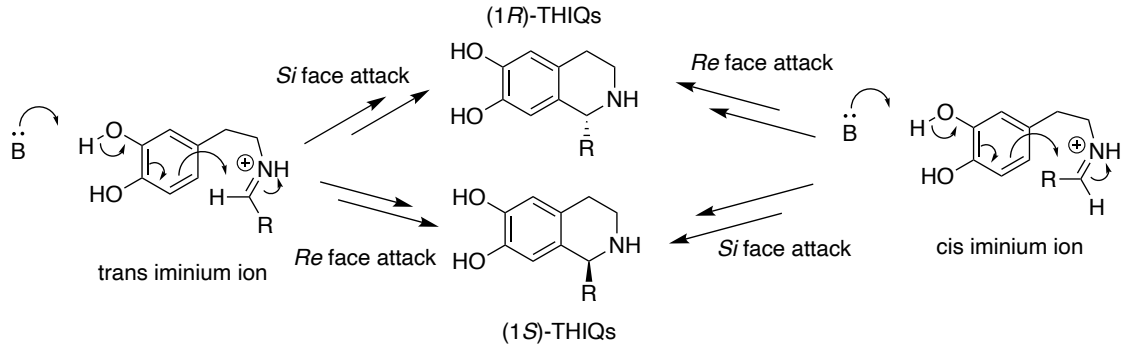


Figure 2.12: The active site entrance/exit of *Tf*NCS. The structure shown is PDB:5NON; co-crystallised structure of wild-type Δ33*Tf*NCS with a non-productive reaction intermediate analogue (**16**) of the reaction between the two natural substrates, **5** and **6**. There are two pockets, one which is accessible to the substrates, and another that is inaccessible, as highlighted in green.

Mechanistically supported by the NCS crystallographic studies, the iminium ion intermediate of the NCS reaction is held in the active site, with the nitrogen of the iminium ion hydrogen-bonded to charged residues E110 and D141. The hydroxyl groups of dopamine interact with the nitrogen of K122. These interactions mean that the aldehyde side chain is forced forwards, forming a trans iminium species. This means that nucleophilic attack from the catechol ring onto the iminium ion can only occur onto the *Re* face of the iminium ion, forming (1*S*)-THIQs (Scheme 2.4). Therefore, if the face of attack or the geometry of the iminium ion can be reversed, the stereochemical outcome of the reaction would be switched.



Scheme 2.4: The stereochemical outcome of a Pictet-Spengler reaction, based upon which face of the *cis* / *trans* iminium ion formed is attacked.

Therefore, to see if the other active site entrance/exit pocket could be accessed, mutations were considered at P179, to glycine, alanine, and serine, to see if this would create extra space. This could allow the aldehyde side chain to sit in either pocket, forming either a *cis* or *trans* iminium ion, altering the stereochemistry of the reaction, but generating racemic *para*-THIQs. It would also be interesting if mutations at this position were tolerated because this proline is at the end of an alpha helix (Figure 2.11) and is known as a 'helix-breaker' residue.[146] Mutations at F80 to more bulky side chains, tyrosine and tryptophan, were considered to increase the hydrophobicity of the active site and to see if the secondary cleft could be blocked further to help increase activity of the enzyme. This was because with some substrates screened, particularly the benzaldehydes, a racemic reaction associated with the enzyme sample had been observed.

These five *Tf*NCS variants, P179A/G/S and F80Y/W were purchased as synthetic genes, codon optimised for *E. coli* expression, and subcloned into pET29a(+). The Δ33*Tf*NCS construct was used, containing an *N*-terminal His-tag and a TEV-cleavage site in case any of the variants were particularly promising and structural studies could be performed by protein crystallography. All five variants were well-expressed, at 17 – 22% of the total protein concentration as determined by SDS-PAGE analysis (Figure 2.13) using ImageJ and Bradford assay.

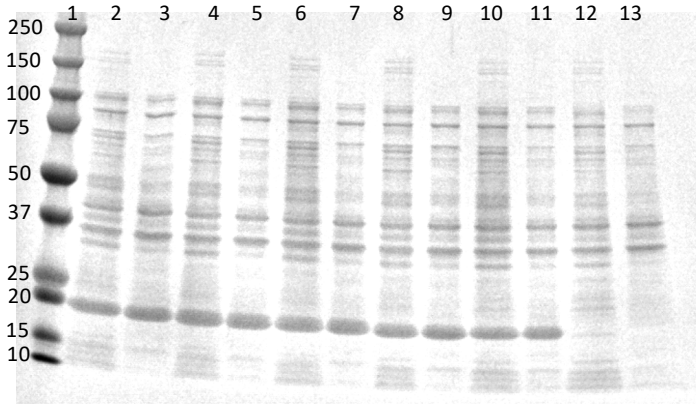


Figure 2.13: SDS-PAGE analysis of Δ33*Tf*NCS variant expression. Expected protein size = 20.7 kDa. Lanes: 1, BioRad Precision Plus Protein™ Dual Color Standards with masses given in kDa. 2, Δ33*Tf*NCS-P179A total protein. 3, Δ33*Tf*NCS-P179A soluble protein. 4, Δ33*Tf*NCS-P179G total protein. 5, Δ33*Tf*NCS-P179G soluble protein. 6, Δ33*Tf*NCS-P179S total protein. 7, Δ33*Tf*NCS-P179S soluble protein. 8, Δ33*Tf*NCS-F80W total protein. 9, Δ33*Tf*NCS-F80W soluble protein. 10, Δ33*Tf*NCS-F80Y total protein. 11, Δ33*Tf*NCS-F80Y soluble protein. 12, empty pET29(+) total protein. 13, empty pET29(+) soluble protein.

Enzymatic assays were performed for the reaction between dopamine (**5**) and phenylacetaldehyde (**76**) or hexanal (**78**) (like Scheme 2.3), with these variants as cell lysate at 0.1 mg mL^{-1} *Tf*NCS. Lysed BL21(DE3) cells containing the empty pET29a(+) vector was used as a control (empty pET, Figure 2.13) and reactions were also performed a no enzyme control (i.e. buffer added rather than purified protein) to compare with the empty pET vector control to and determine the origin of the background Pictet-Spengler reaction. Comparable conversions were observed for both reactions (Figure 2.13) so the background reaction is associated with the reaction buffer, rather than any phosphates or other proteins present in the cell lysate. All five *Tf*NCS variants were highly stereoselective with e.e. >92% in the product for reactions with **76** and e.e.s >98% for reactions with **78**. A significant background reaction is observed (generating around 30% racemic product) so it can be assumed that any (*R*)-product generated is by this reaction and so reflects a decrease in enzyme activity. **76** is more activated than **78**, which likely accounts for the generally lower

stereoselectivities observed. Conversions and enantiomeric ratios of the products generated are given in Figure 2.14. P179S had comparable activity with the wild-type enzyme and variants, F80W and P179A gave slightly improved conversions. The variant P179G gave decreased conversions for reactions with phenylacetaldehyde but comparable activity to the wild type with hexanal, perhaps reflecting a structural change in the exit of the active site, hindering the access of the larger, benzylic aldehyde. F80Y was also less active, perhaps due to the hydroxyl group of tyrosine reducing favourable hydrophobic interactions which promote substrate binding. It is therefore clear that the mutations P179A/G/S alone cannot alter the NCS reaction selectivity. It is unclear whether these variants create enough space to allow the aldehyde to access the secondary cleft of the active site entrance/exit nor the effect of the variants F80W/Y. Therefore, future work could explore double point variants at positions 80 and 179 or structural investigations. Future rational design attempts could focus on the charge residues E110 and D141 to hold the iminium ion intermediate on the other side of the active site, to hopefully either generate the *cis* iminium ion or force *Si* face attack in the cyclisation reaction step (Scheme 2.4).

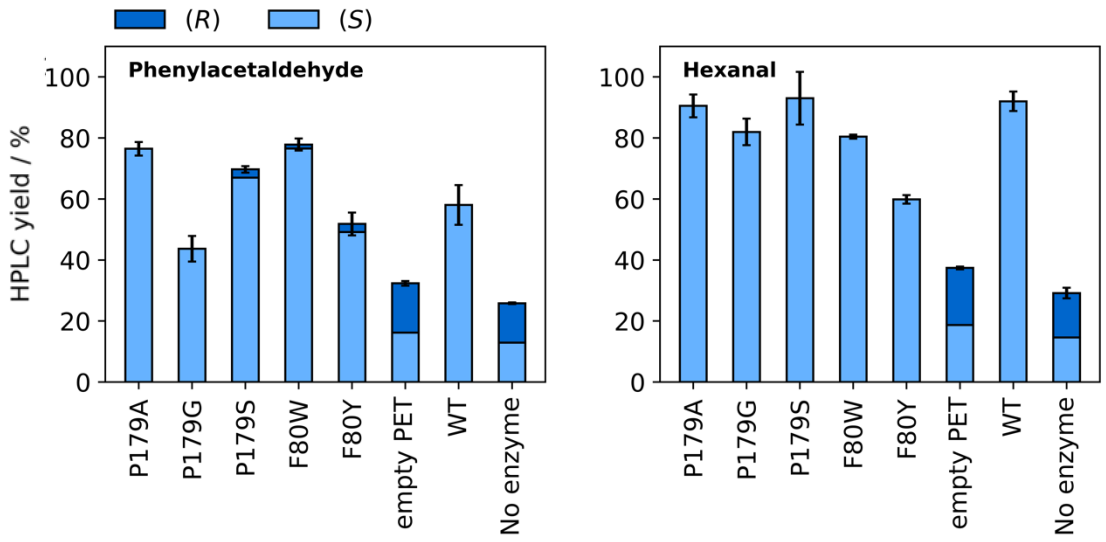
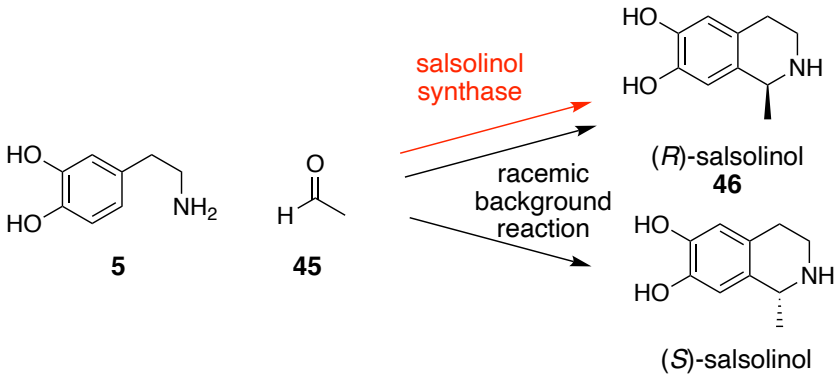


Figure 2.14: Conversions of reactions using single point variants of Δ33*Tf*NCS at positions 179 and 80. Reactions were performed using 0.1 mg mL^{-1} enzyme under standard reaction conditions for reactions between **5** and phenylacetaldehyde (**76**) or hexanal (**78**). Conversions were determined by product formation calibration curves (see Appendix) and

enantiomeric excesses determined by chiral HPLC (method 5) which are given above each bar and enantiomeric ratios are shown by split bars.

2.2.8 Salsolinol synthase

Another biocatalytic route towards (1*R*)-THIQs may already exist in nature. Salsolinol synthase, a ubiquitin-like enzyme found in rats, is thought to catalyse a stereoselective Pictet-Spengler reaction between dopamine and acetaldehyde, giving (*R*)-salsolinol (**56**) (Scheme 2.5). The enzymatic reaction is thought to compete with a racemic background Pictet-Spengler reaction as both enantiomers of salsolinol are isolated, with the (*R*)-enantiomer in excess. However, isolation of the other possible THIQ regioisomer, the *ortho* product has not been reported which would also be formed if a racemic Pictet-Spengler reaction is present.[15]



Scheme 2.5: The formation of salsolinol in rat brain. A racemic Pictet-Spengler reaction competes alongside a stereoselective enzymatic reaction by the putative enzyme, salsolinol synthase.

Increased levels of expression of the enzyme in the brain have been shown to be correlated with increased production of **46** in both humans and rats, although the molecular weights of the isolated proteins differ greatly (*ca*. 33 kDa vs. 8.6 kDa). The sequence is only known from the variant isolated from *Rattus norvegicus* and is a ubiquitin-like protein, consisting of 77 amino acids. The sequence of ubiquitin is highly conserved throughout nature. Salsolinol synthase differs by only four residues and is the only known enzymatic homologue of ubiquitin.[107] The recombinant expression and purification of the enzyme has already been reported but only a limited investigation of

the catalytic scope has been performed and for the enzyme *in vivo* rather than using the isolated protein.

Three different constructs of salsolinol synthase were purchased as synthetic genes, codon optimised for *E. coli* expression, each with different promotors, pTAc, pP1 or T7 for salsolinol synthase SSv1, SSv2 and SSv2 respectively and all were transformed into BL21(DE3). *C*-terminal His-tags were also introduced to aid protein purification if required, although the protein has been previously isolated by acid precipitation as it is reported to be highly soluble.[147]

Expression was performed using analogous conditions to those used for *Tf*NCS expression. Ubiquitin is highly stable and high levels of protein can be generated from recombinant expression; therefore, it was hoped that expression conditions need not be optimised. Unfortunately, only a small band around 10 kDa was observed by SDS-analysis of SSv1 expression and no expression was observed for the other two constructs (Figure 2.15).

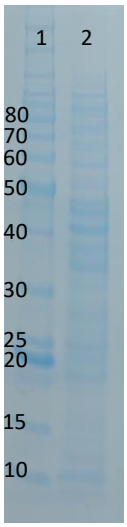


Figure 2.15: **SDS-PAGE analysis of salsolinol synthase expression.** Lanes: 1, Benchmark™ Protein Ladder. 2, supernatant of cell pellet lysis after salsolinol synthase expression. Protein expected at 9.7 kDa.

To determine whether SSv1 had in fact been expressed, purification was attempted following the reported protocol. Due to the apparent high stability of the enzyme, it can be purified via acid precipitation of the other proteins. Removal of the other, precipitated proteins by centrifugation and SDS-PAGE analysis of the supernatant showed that protein was present, however despite the use of a reducing agent, oligomers were observed. To check if the isolated protein was active, enzymatic reactions were performed between dopamine and acetaldehyde, however no THIQ products (**46**) were observed above the background reaction as determined by RP-HPLC and LCMS.

The poor or no expression levels observed suggested that there were issues with the constructs designed. Since salsolinol synthase has such a short sequence, perhaps the introduction of a terminal His-tag affects the stability of the enzyme. Therefore, in future investigations, the His-tag should be omitted. Due to the stability of the protein, isolation of the pure protein theoretically can be achieved by acid precipitation and so tags are not required.

2.3 Conclusions and future work

To conclude, crystallographic investigations of *Tf*NCS were explored to try and further understanding of the mechanism and substrate acceptance. The previously published truncated construct, Δ33*Tf*NCS was used, however a further truncation on the *N*-terminus, removing residues 1-42 rather than 1-33, resulted in no protein expression. Reaction intermediate mimics based on reactions with aliphatic aldehydes were synthesised and co-crystallisation was attempted. However, the molecules were not observed clearly in the active site, likely due to conformational flexibility and a lack of electron density. In future, mimics based upon reactions with aldehydes containing an aromatic ring system should be considered as these have proven promising.[30,148,149] Structural investigations by NMR spectroscopy could also be considered; previous studies have shown that NCS has a highly dynamic mechanism and a lot more is now

known about the mechanism and substrate scope. Protein NMR may therefore be a useful tool to fill in the gaps in our knowledge.

Routes towards generating (1*R*)-THIQs were also explored: the attempted expression and purification of several enzymes, salsolinol synthase and two NCSs from the sacred lotus and, single point variants of *Tf*NCS. Although little change in stereoselectivity was observed where enzymes were expressed and active, these investigations provide a useful starting point for future studies. Different constructs of salsolinol synthase and the *Nn*NCSs (such as removing His-tags, changing terminus of His-tags or truncating termini) would likely improve expression levels and allow for screening. Future mutagenesis of *Tf*NCS could also be explored by rational design or directed evolution.

Chapter 3: The Acceptance and Kinetic Resolution of α-Methyl Substituted Aldehydes by Norcoclaurine Synthases

3.1 Introduction

3.1.1 The tetrahydroisoquinoline moiety as a privileged drug scaffold

The THIQ moiety is found in many pharmaceutically relevant molecules, so is known as a privileged drug scaffold. This is a molecular moiety that has already been shown to be present in ligands for a range of different receptors, making them an attractive target in drug discovery efforts.[150]

The THIQ scaffold is found in both naturally occurring and synthetic drug molecules, with a vast range of different biological activities exhibited. In plants, many BIAs contain the THIQ moiety and have been shown to have useful medicinal properties, including anti-microbial (sanguinarine (**12**) and berberine (**2**)), anti-tussive (glaucine (**80**)) and anti-cancer (noscapine (**3**)).[1] Other naturally-occurring THIQs with biological activities include saframycin, a non-ribosomal peptide with anti-microbial activities and trolline (**58**), an alkaloid with anti-viral properties.[119,151] Useful synthetic THIQs include solifenacin (**81**), a widely prescribed anti-diuretic, almorexant (**82**), with anti-insomniac properties and quinapril (**83**), an ACE inhibitor used in the treatment of hypertension.[55,152,153] The chemical structures of a range of pharmaceutically relevant THIQs are given in Figure 3.1.

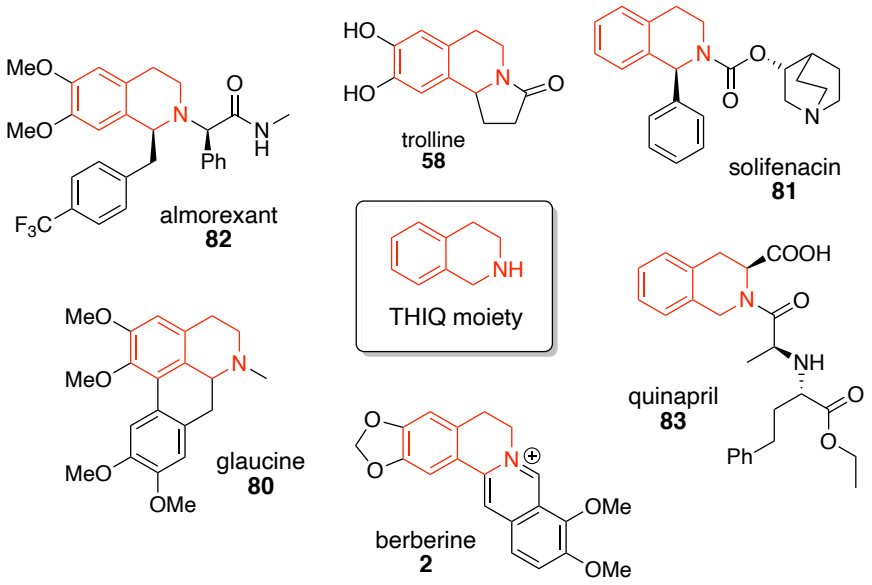


Figure 3.1: A range of biologically relevant molecules containing the THIQ moiety (given in red).

3.1.2 Synthetic routes towards the THIQ scaffold

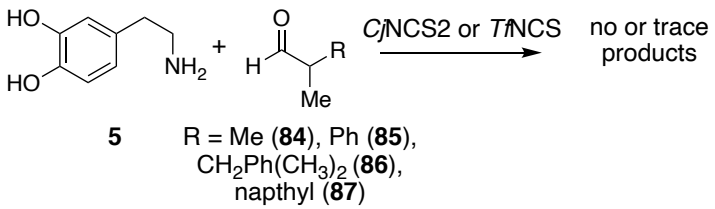
The pharmaceutical relevance of the THIQ moiety has led to the development of a range of synthetic routes for its formation, as discussed in-depth in Section 1.6. Although there is no generally accepted best route, common methods employ Bischler-Napieralski or Pictet-Spengler reactions.[62,154] Gaining regio- and stereoselective control often requires a multi-step strategy, metal catalysts (which can be polluting, expensive or a limited resource) or toxic reaction conditions.[15,76,155] To develop more facile syntheses towards single isomer THIQs, the use of suitable biocatalysts has been explored. Enzymes are advantageous compared with traditional synthetic methods as they often exhibit high regioselective and stereoselective control, and operate under mild reaction conditions, however, can be limited by a small substrate scope.

Several enzymes have been used for the biocatalytic syntheses of chiral THIQs, including monoamine oxidases (MAOs) and imine reductases (IRs), however there is the synthetic requirement to generate the racemic THIQ or dihydroisoquinoline precursor, respectively.[67,156,157] The use of Pictet-Spenglerase enzymes to generate the scaffold avoids this synthetic requirement, forming the THIQ scaffold in a single, regio-

and stereoselective step. There are a few known Pictet-Spenglerase enzymes capable of generating a THIQ scaffold, as discussed in Section 1.7, however by far the most widely explored is NCS, involved in BIA biosynthesis.[19] The NCS enzyme isolated from *Thalictrum flavum*, *Tf*NCS has been shown to have a substrate scope that is much wider than the natural substrate scope, accepting a range of dopamine analogues as the amine substrate and various aliphatic and benzylic aldehydes or ketones as the carbonyl substrate, thus generating a wide range of pharmaceutically-relevant THIQs.[21,36,118]

3.1.3 Alpha-substituted aldehydes as NCS substrates

Previous investigations into the NCS substrate scope have reported poor tolerance towards α-substituted aldehydes (**84**-**87**, Scheme 3.1). Previously, when Ruff *et al.* explored the use of Δ29*Tf*NCS with 2-(1-napthyl)propanal (**87**) and cyclic carboxaldehydes with dopamine, no Pictet-Spenglerase products were detected, which was rationalised as being due to a lack of conformational space in the active site.[21,31] The acceptance of 1-naphthalene acetaldehyde but not 2-(1-napthyl)propanal (**87**) also led to the rationale that more generally α-methyl substituted aldehydes were not well tolerated by NCS (Scheme 3.1). Interestingly, Nishihachijo *et al.* and Pesnot *et al.* observed low levels of dopamine consumption when using α-methyl substituted aldehydes isobutanal (**84**) and 2-phenylpropanal (**85**) with *Coptis japonica* NCS (*Cj*NCS2), indicating that an NCS catalysed reaction may have occurred but no products were isolated or characterised.[22,158]



Scheme 3.1: NCS α-substituted aldehyde substrate scope demonstrated in previous work. This has been described by Ruff *et al*.[21], Nishihachijo *et al*.[35] and Pesnot *et al*.[22]

Incorporating α-methyl aldehydes into THIQs is challenging using traditional organic synthetic methods as potentially four different isomers can be formed if the reaction is regioselective and eight different isomers if not. If a single enantiomer of aldehyde was used, which two diastereomers would be generated. Many single-isomer α-methyl substituted aldehydes are not commercially available and have been accessed via another available chiral molecule or synthesised using chiral auxiliary methods[159,160] and via asymmetric hydroformylation[161]. However, these methods are step-intensive, often involve the use of toxic reagents and gaining a high e.e. in the product is challenging due to racemisation. The resulting aldehyde products can also be volatile and prone to oxidation.[162] Thus, incorporating a single enantiomer aldehyde into the product using non-enzymatic routes is non-trivial and can lead to poor enantiopurity in the product.

Accessing the resulting THIQs from reactions between a dopamine analogue and an α-methyl aldehyde is also pharmaceutically relevant. Products of this reaction have been reported to possess bronchodilatory activities (**88**) and a corresponding isoquinoline (**89**) has also been shown to have antimalarial activities.[3,163] Indeed, corydaline (**90**), a 13-methyltetrahydroprotoberberine alkaloid is also a major constituent of *Corydalis tuber* which is in clinical trials in Korea for the treatment of functional dyspepsia.[164] The structures of these compounds are given in Figure 3.2.

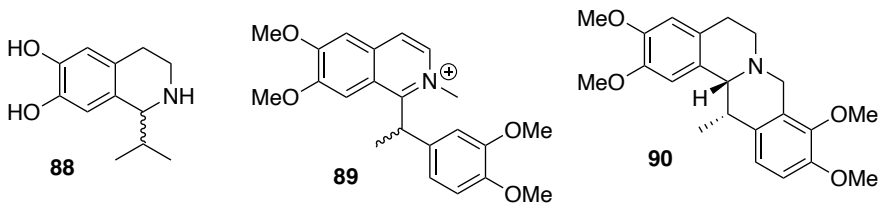
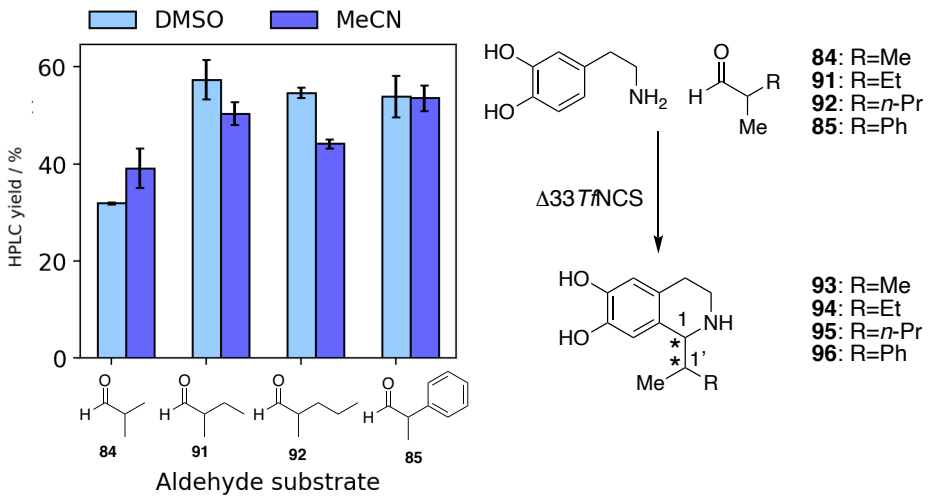


Figure 3.2: Structures of pharmaceutically relevant THIQs based upon Pictet-Spengler reactions involving an α-methyl substituted aldehyde as the carbonyl substrate.

3.2 Results and Discussion

3.2.1 Initial NCS reactions

Following the reported acceptance of ketones as substrates,[21] with a view to significantly extending the scope of the Pictet-Spenglerase NCS towards more sterically challenging aldehydes, a range of α-methyl and α-ethyl substituted aldehydes were initially tested as substrates. Reactions were first performed using 1 equivalent of dopamine (**5**) and 1 equivalent of aldehyde (**84, 85, 91, 92**) (Scheme 3.2) using either DMSO or acetonitrile as a co-solvent as previous studies have shown that the optimal co-solvent can be substrate dependent.[31] For NCS-catalysed reactions with ketones, using DMSO as the co-solvent gives superior reaction rates compared with using acetonitrile, with five-fold increases in product formation observed. This is hypothesised to be due to shifting the reaction equilibrium from amine/ketone towards imine formation, thus enhancing enzymatic turnover.[31] This has also been shown to be the case with other enzymes, with the presence of different organic solvents capable of affecting the stability, reaction rate and selectivity.[165] Sodium ascorbate (1 e.q.) was added to the reaction mixture to reduce oxidation of **5** and the aldehyde substrates, both of which are highly sensitive to oxidation.[20]



Scheme 3.2: **NCS reactions initially performed between dopamine and a range of α-methyl substituted aldehydes**. *Reaction conditions with NCS*: dopamine HCl (10 mM), aldehyde (10 mM), sodium ascorbate (5 mM), Δ33*Tf*NCS (0.2 mg mL^{-1}) in HEPES buffer (100 mM, pH 7.5) with DMSO (10% *v/v*) at 37 °C for 3 h, 250 μL scale reactions. Conversions determined based on product formation compared to calibration curves of the products.

Analytical HPLC results (method 1) showed the presence of product peaks (**93** – **96**) for all aldehydes tested in the expected range (6-8 min).[31] Reactions were therefore performed on a preparative scale to enable conversions to be determined based upon calibration curves of purified product standards. Dopamine and aldehydes are both oxidatively sensitive so conversions could not be determined by starting material depletion.

Conversions were shown to be 32-57% (Scheme 3.2), determined by measuring product formation using HPLC analysis (method 1), compared to calibration curves of purified products. Using DMSO as the co-solvent resulted in slightly improved conversions with two of the aldehydes tested (**91** and **92**), therefore in later enzymatic reactions, this co-solvent was used rather than acetonitrile. It is unknown why a more significant solvent-effect is observed with ketones rather than with α-methyl substituted aldehydes. However, iminium ion formation is more readily favoured with aldehydes compared with ketones, as aldehydes are more electrophilic, so the solvent used likely

has less of an effect on iminium formation. The conversions observed with these substrates were also significantly higher than those reported previously. Nishihachijo *et al*. reported only 6% and 8% relative activity for reactions between dopamine **5** and aldehydes **84** or **85**, respectively using *Cj*NCS2 (NCS isolated from *Coptis japonica*) as the biocatalyst.[158] The reaction conditions reported were using 10 mM of both substrates with 1% DMSO as the co-solvent for 15 min. Lower conversions could therefore be accounted for by shorter reaction times (15 min vs. 3 h), lack of reducing agent to stabilise the dopamine, issues of aldehyde solubility and differing substrate scopes of *Cj*NCS2 and *Tf*NCS.

*Cj*NCS2 has previously been shown to have a different substrate scope to *Tf*NCS, with ketones not being accepted as substrates. No product was observed for *Tf*NCS-catalysed reactions between dopamine (**5**) and 2-naphthalen-1-ylpropanal (**87**) (Scheme 3.1) in reactions performed by Ruff *et al*.[21] However, this aldehyde is significantly bulkier than **85**, so it is possible that it does not fit into the enzyme active site. Lower substrate loading was also used (1 mM amine and 1 mM aldehyde) by Ruff *et al.* and recent studies have shown that increased substrate loading (*ca*. 10 mM) improves conversions, particularly with non-natural substrates.[36] This is unsurprising considering that the K_d of *Tf*NCS for dopamine (**5**) is 5 mM.[26]

^{1}H-NMR analysis of the products generated using a racemic aldehyde substrate (**85**, **91**, **92,** Scheme 3.2) showed two different environments for the 1-H proton (4-5 ppm), with a major and minor product formed, thus suggesting that two diastereomers of the product were present (Figure 3.3). *Tf*NCS has been reported previously to generate THIQs with *S*-stereochemistry at the C-1 position and it was rationalised that this stereochemical preference is retained. As a racemic aldehyde was used, it is possible that only one enantiomer of aldehyde was accepted, i.e., a kinetic resolution of the aldehyde occurred. Therefore, if equal molar quantities of amine and aldehyde were

used, the maximum conversion possible would be 50%, if no *in situ* racemisation of the aldehyde takes place and that only one enantiomer of aldehyde is accepted. One equivalent of aldehyde was used for initial screening with α-methyl substituted aldehydes (Scheme 3.2) and the maximum conversion observed in the reactions was 57%. A conversion >50% may also indicate that some aldehyde racemisation occurred.

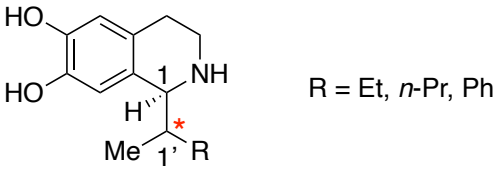
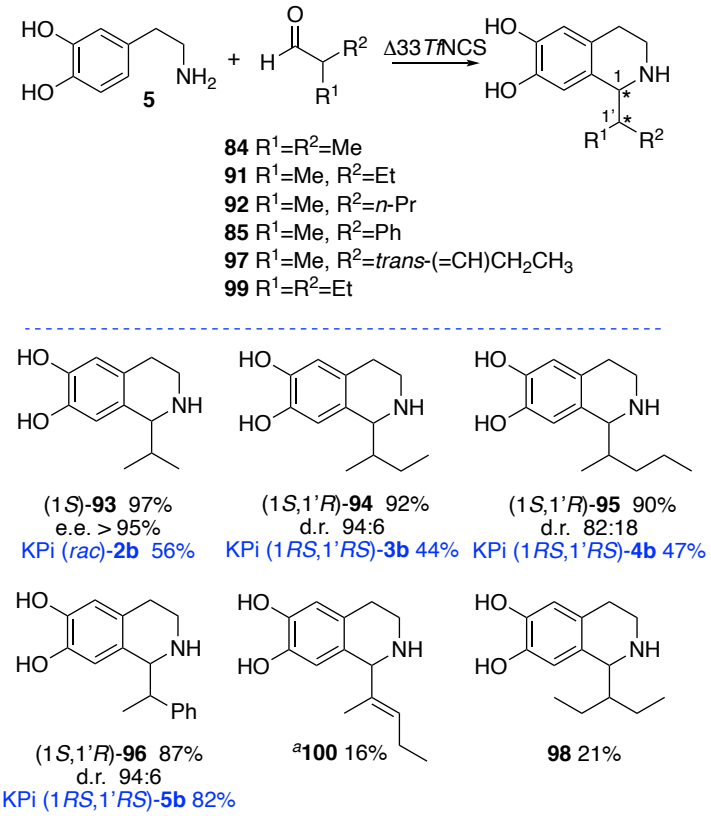


Figure 3.3: Two chiral centres are generated in the THIQs resulting from NCS reactions between dopamine and racemic α-methyl substituted aldehydes.

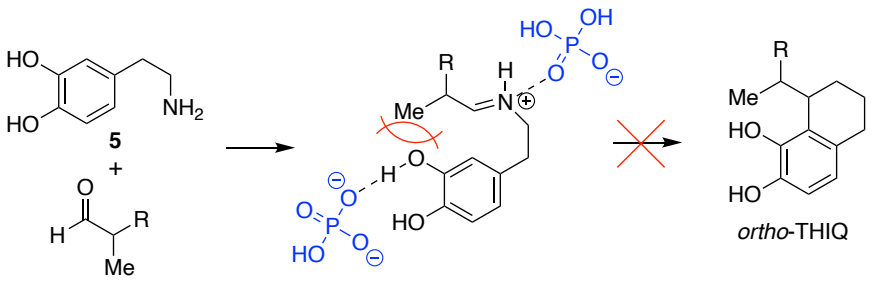
Reactions using two equivalents of aldehyde were investigated using wild-type NCS (WT-Δ33*Tf*NCS) and dopamine **5** with aldehydes **84**, **85**, **91**, **92** (Scheme 3.3), using otherwise the same conditions as described in Scheme 3.2. Conversions were improved, with all α-methyl substituted aldehydes (**84**, **85**, **91**, **92**) readily accepted by *Tf*NCS to give products **93-96** in high HPLC yields (87-97%). Yields decreased slightly with increased steric bulk of the aldehyde, probably due to steric constraints in the enzyme active site. Aldehyde **97** was probably less readily accepted likely due to less rotational freedom of the alkyl chain in the active site and reduced reactivity of the conjugated aldehyde. THIQ **98,** formed by reaction with an α-ethyl-substituted aldehyde **99**, was also investigated to see whether the substrate scope could be further extended, but the product was formed in a significantly lower HPLC yield (21%) most likely due to steric constraints.



Scheme 3.3: *Tf*NCS reactions with dopamine and the α-methyl and α-ethyl substituted aldehydes using two equivalents of aldehyde. *Reaction conditions with NCS (black)*: dopamine.HCl (10 mM), aldehyde (20 mM), sodium ascorbate (5 mM), Δ33*Tf*NCS (0.2 mg mL^{-1}) in HEPES buffer (100 mM, pH 7.5) with DMSO (10% *v/v*) at 37 °C, 250 μL scale reactions. *Reaction conditions with KP (blue)i*: dopamine.HCl (10 mM), aldehyde (20 mM), ascorbic acid (5 mM) in potassium phosphate buffer (KPi) (0.3 M, pH 6) with acetonitrile (50% *v/v*) at 60 °C. Yields were determined by monitoring product formation against standards (by HPLC) unless otherwise indicated; diastereomeric ratios (d.r.s) for stereoselective reactions were determined by ^{1}H-NMR spectroscopy[a]. The conversion was determined by monitoring dopamine consumption for products **98** and **100**. Figure reproduced from Roddan *et al.*[148]

Analogous reactions to the higher yielding enzyme reactions were also performed without enzyme using potassium phosphate (KPi) and acetonitrile (1:1) to give the racemic products **93-96** in 44-82% yield (given in blue, Scheme 3.3).[15] Comparison of the RP-HPLC data with that generated by an NCS-mediated reaction indicated that no or traces of the *ortho*-product were formed in all cases, reflecting the more sterically demanding nature of the α-substituted aldehydes.[15] For reactions with linear, aliphatic, or benzylic aldehydes, some background, racemic Pictet-Spengler reaction is often

observed.[31] No reactions were observed using **97** and **99** as substrates, likely due to steric constraints in the reaction transition state where there is an unfavourable interaction between the aldehyde side-chain and the *meta*-hydroxyl of dopamine (**5**), as demonstrated in Scheme 3.4.



Scheme 3.4: Trace levels of *ortho*-THIQs were formed during KPi-mediated Pictet-Spengler reactions between dopamine and α-methyl aldehydes. This is likely to be due to steric clashes between the *meta*-hydroxyl group and the methyl or R groups of the aldehyde.

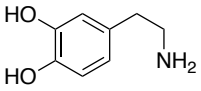
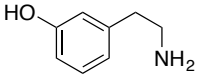
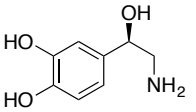
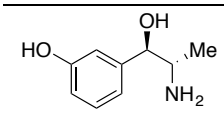
In all cases, reaction yields were lower in reactions using KPi as a catalyst (no NCS present) compared to with NCS (Scheme 3.3). This also provided a rationale as to why no background reaction was observed in NCS-catalysed reactions when replacing the protein sample with an equal volume buffer. The racemic product (*rac*-**93**) generated using the phosphate-mediated Pictet-Spengler reaction and isolated for use as a racemic standard for the chiral HPLC analysis for NCS-generated **93**. The increased reactivity of NCS compared with KPi as the reaction catalyst highlights the synthetic value of the NCS-mediated reactions. The phosphate-mediated reaction with these substrates is however useful for generating racemic products in a regioselective manner.

3.2.2 Extension of the amine substrate scope

To extend the range of THIQ products accessible by an NCS reaction involving α-methyl substituted aldehydes, a range of other amine analogues were screened. Reactions were also observed with *meta*-tyramine (**51**), norepinephrine (**101**) and metaraminol (**102**). Conversions were based on amine consumption from RP-HPLC

based calibration curves, as these amines explored are less oxidatively sensitive than dopamine (**5**). The conversions gained are given in Table 3.1. In all cases, conversions of these amines (**51**, **101** and **102**) using aldehydes (**84**, **85**, **91**, **92**) were lower than those with dopamine (**5**). The products were therefore not isolated and instead product formation was confirmed by LC-MS. The lower conversions are most likely due to the different binding modes of the unnatural dopamine substrates into the active site leading to a decreased acceptance of the more challenging aldehyde substrates. Slightly increased conversions were observed for reactions between **51** and aldehydes **97** and **99** compared to the analogous reactions using dopamine (**5**) as the amine substrate. Despite lower conversions generally, the expansion of the amine substrate scope with these aldehydes highlights the more general applicability of using α-methyl substituted aldehydes as NCS substrates.

Table 3.1: Conversions when using α-substituted aldehydes and a variety of different amine analogues. *Reaction conditions*: amine (10 mM), aldehyde (20 mM), sodium ascorbate (5 mM), Δ33*Tf*NCS (0.2 mg mL^{-1}) in HEPES buffer (100 mM, pH 7.5) with DMSO (10% *v/v*) at 37 °C for 18 h, 250 μL scale reactions. ✓= product observed but no conversion determined. x = no conversion (i.e. no product observed). Reactions were performed in triplicate and were consistent within 10% error. Conversions determined based on amine consumption (HPLC method 1) and confirmed by the detection of the products by LC-MS. Figure adapted from Roddan *et al*.[148]

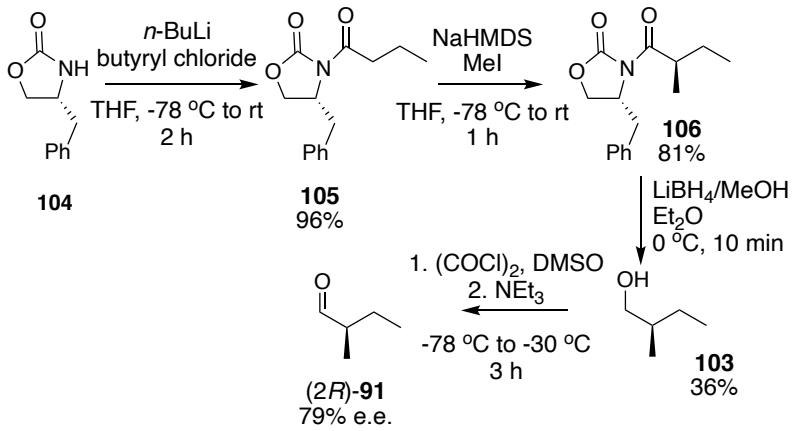
Amine	Aldehyde					
	84	**91**	**92**	**85**	**97**	**99**
 **5**	97%	92%	90%	87%	16%	21%
 **51**	34%	48%	45%	69%	25%	26%
 **101**	✓	65%	57%	79%	x	x
 **102**	64%	50%	42%	69%	x	x

3.2.3 Determination of the reaction stereoselectivity

For reactions with α-methylated aldehydes, stereoselectivities were explored for the highest yielding reactions with dopamine (**5**). Aldehyde **84** is achiral and chiral HPLC analysis of the product **93** indicated it was formed in >95% e.e., by comparison to the racemic standard prepared using the KPi-mediated reaction (Scheme 3.3). This was assigned as the (*S*)-isomer, based upon literature precedent where NCS has been shown to exclusively generate THIQs with (*S*) stereochemistry at C-1. The elution order of the two enantiomers during chiral HPLC analysis using a Astec® Chirobiotic T2 column was also consistent with the literature, whereby the (*S*)-enantiomer of THIQs always elutes first.[14,22] THIQs **94**-**96** possess two chiral centres, and analysis of the ^{1}H NMR spectra revealed two proton environments for 1-H (*ca*. 4-5 ppm), indicating that two diastereomers of product had been formed. A significant majority of one diastereomer was observed. Diastereoisomeric ratios (d.r.s) were determined by ^{1}H NMR spectroscopy, and high d.r.s were noted for **94** and **96** (94:6) (Scheme 3.3). A lower diastereomeric ratio was observed in **95** (82:18). Since wild-type *Tf*NCS typically generates (*S*)-THIQs at C-1, it was rationalised that the stereoselectivity at this position was retained and that a kinetic resolution of the aldehyde had occurred leading to a mixture of (*S*) and (*R*) stereochemistry at the C-1' position.

To determine which aldehyde was preferentially accepted and thus determine the predominate stereochemistry at the C-1' position, the synthesis of two single enantiomer aldehydes, (2*R*)-**91** and (2*S*)-**91** was explored. Aldehyde **85** was not prepared due to the fast racemisation rates reported at pH 7.5, and aldehyde **91** was chosen over **92** because higher d.r.s in the resultant THIQ product were observed.[166] The analogous alcohol to aldehyde **91**, (*S*)-2-methyl-butan-1-ol ((*S*)-**103**), is commercially available in high e.e. (>95%) and the synthesis of (*R*)-2-methyl-butan-1-

ol ((2*R*)-**91**)was investigated via the use of an Evan's auxiliary, as shown in Scheme 3.5.



Scheme 3.5: **Synthetic route towards (2*R*)-91.**

The first step, the amide coupling of the oxazolidinone (**104**) to butyryl chloride, proceeded to give the desired product, **105** in high yield (96%) after flash column chromatography.[167] To achieve high selectivity in the methylation step, it was found after several attempts that the methyl iodide needed to be added slowly, over five minutes, and cooled to 0 °C before addition to the reaction mixture. A high diastereomeric ratio in the methylated product (**106**, 95:5) was achieved with only trace amounts of the other diastereomer (with *S* stereochemistry in the chiral centre generated after methylation) present.[167] The reaction to remove the oxazolidinone group to give the alcohol used literature reaction conditions, however isolation of the alcohol was challenging. Initially, the reaction was performed using $LiAlH_4$, however the alcohol was trapped by the aluminium salts and could not be isolated by extraction or distillation. The reaction was therefore attempted using lithium borohydride in MeOH.[168] After a straightforward workup (addition of $NaHCO_3$ followed by extraction into diethyl ether), a mixture of the alcohol (**103**) and oxazolidinone (**104**) remained. Attempts to isolate the alcohol by distillation were unsuccessful, likely due to trapping by the solid oxazolidinone and the high volatility of the alcohol. Therefore, the alcohol was obtained by column chromatography using low boiling point solvents, diethyl ether and *n*-hexane, as the mobile phase. Oxidation of the alcohol (2*R*)-**103** to (2*R*)-**91** was initially

attempted using Dess-Martin periodinane, however separation of the aldehyde from the Dess-Martin reagent was not possible by extractive workup as both aldehyde and Dess-Martin periodinane are organic and water soluble. The aldehyde is also highly volatile and prone to racemisation and polymerisation, so a simple extraction process was required without any further purification steps. Therefore, a Swern oxidation was used to give the final product, (2*R*)-**91**.[169] The resulting residue was partially purified by silica chromatography immediately after workup then used directly for enzymatic reactions without further purification to avoid issues of aldehyde volatility, racemisation and polymerisation. The aldehyde, (2*S*)-**91** was prepared by an analogous Swern oxidation from the commercially available alcohol, (2*S*)-**103** (commercially available at 95% e.e.) and was obtained in 68% e.e. ^{1}H-NMR spectroscopic data indicated reasonable levels of purity for both aldehydes, and chiral GC analysis was used to determine e.e.s (with the method devised by Helena Philpott, Hailes group). Chiral GC analyses of the two aldehydes compared with a racemic standard are given in Figure 3.4.

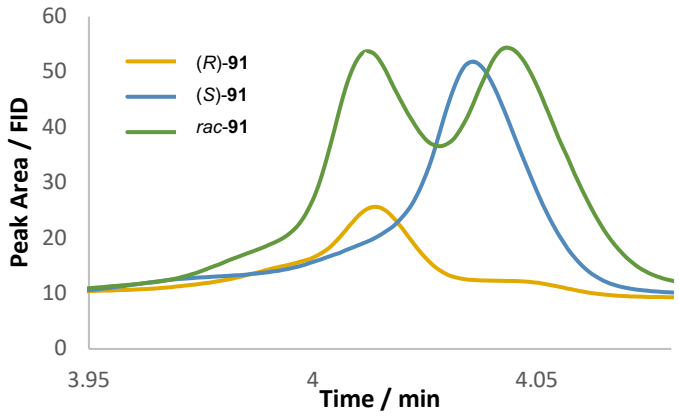
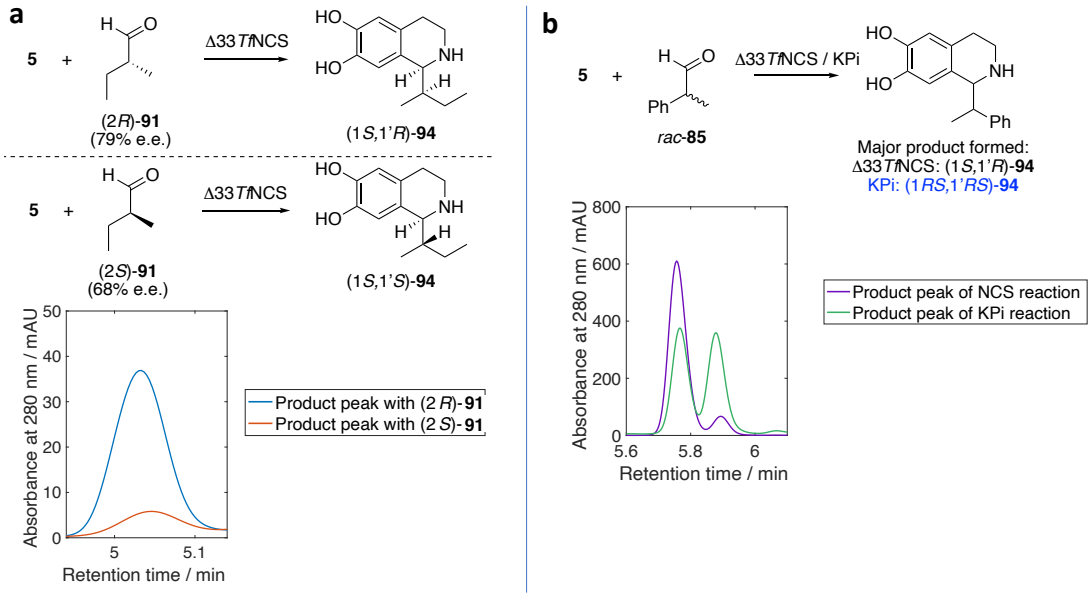


Figure 3.4: Chiral GC analysis of (2*R*)-91 and (2*S*)-91. Figure adapted from Roddan *et al*.[148]

*Tf*NCS reactions were performed with both (2*R*)-**91** and (2*S*)-**91** using the wild-type enzyme using the same reaction conditions as previously described, except 3 e.q. of aldehyde were added to account for impurities present. Yields were significantly lower than previously reported (Scheme 3.3), most likely due to aldehyde impurities present. HPLC analysis (Scheme 3.6a) of the reactions under the same conditions indicated that (2*R*)-**91** was accepted preferentially to the (2*S*)-isomer, confirming that a kinetic

resolution was occurring when using the racemate. Indeed, the low levels of product formation with (2*S*)-**91** (68% e.e.) most likely arose from the small amount of (2*R*)-isomer present (~16%) or *in situ* racemisation. Following the established selectivity at C-1 for *Tf*NCS and the NMR spectroscopic data, it was concluded that the major isomer formed when using ***rac*-91** was (1*S*,1'*R*)-**94** (d.r. 94:6). It was also rationalised that the same stereoselectivity most likely arises when using **92** to give (1*S*,1'*R*)-**95** (d.r. 82:18) as the major isomer out of the 4 possible isomers (Scheme 3.6b).



Scheme 3.6: Analytical analysis of the stereochemical outcomes of NCS reactions with α-methyl substituted aldehydes. a. Analytical HPLC analysis (method 1) of enzymatic reactions between dopamine (**5**) and isomers (2*R*)-**91** and (2*S*)-**91**. *Reaction conditions:* **5** (10 mM), (2*R*)-**91** (30 mM) or (2*S*)- **9** (30 mM) sodium ascorbate (5 mM), purified Δ33*Tf*NCS (0.2 mg mL^{-1}) in HEPES buffer (100 mM, pH 7.5) with DMSO (10% *v/v*). The total reaction volume was 250 μL and they were performed at 37 °C with results confirmed in triplicate. **b.** Analytical HPLC analysis of enzymatic and KPi reactions between **5** and **85**. *Reaction conditions for enzymatic reaction:* dopamine.HCl (10 mM), **85** (20 mM), sodium ascorbate (5 mM), purified Δ33*Tf*NCS (0.2 mg mL^{-1}) in HEPES buffer (100 mM, pH 7.5) with DMSO (10% *v/v*) at 37 °C for 18 h, 250 μL reaction volume. *Reaction conditions for KPi reaction:* dopamine.HCl (10 mM), **85** (20 mM), sodium ascorbate (20 mM) in KPi buffer (0.3 M, pH 6) with acetonitrile (50% *v/v*) at 60 °C for 18 h, 1 mL reaction volume. Figure adapted from Roddan *et al.*[148]

RP-HPLC analysis of racemic THIQ **96** generated using the non-stereoselective KPi-mediated Pictet-Spengler reaction[15], revealed two product peaks corresponding to 4-isomers (2 sets of diastereoisomers) in a ratio of 1:1 (Scheme 3.6b). ^{1}H-NMR spectroscopy confirmed that only a trace of the *ortho*-product was formed. When compared to HPLC data for the enzymatic reaction, it was evident that an excess of one diastereomer was formed (ratio ~95:5) (Scheme 3b) in the enzymatic reaction which was consistent with the ^{1}H NMR data. Since the diastereomers, but not the enantiomers are separable by RP-HPLC, it was concluded that the first peak (retention time = 5.8 min) corresponded to (1*S*,1'*R*)-**96** and (1*R*,1'*S*)-**96** and the second peak (retention time = 5.9 min) corresponded to (1*S*,1'*S*)-**96** and (1*R*,1'*R*)-**96**. Since arylacetaldehydes react with *Tf*NCS to give the (*S*)-isomer at C-1 in high e.e.s,[22] and following the enantioselectivity data for **91**, this highlighted that the major product was likely to be (1*S*,1'*R*)-**96** (d.r. 94:6). Attempts to further confirm this result by chiral HPLC separation of the four different stereoisomers present were unsuccessful and the diastereomers could not be separated by preparative HPLC. Clearly, the enzymatic reaction with wild-type *Tf*NCS is an excellent route to such THIQs, as the reactions give higher yields than the KPi reaction and there is potentially stereochemical control at both chiral centres.

3.2.3 Scale-up biotransformation

There are often challenges associated with performing biocatalytic reactions at scale. To determine whether the NCS-catalysed reactions involving α-methyl substituted aldehydes could be readily scaled up, a 0.5 g scale reaction was performed using clarified cell lysate of Δ33*Tf*NCS with **5** and **91** as the aldehyde substrate. The product was isolated via an extractive workup (method 7), developed by Jianxiong Zhao (Hailes group) involving the precipitation of the enzyme by addition of acid (1 M HCl) and centrifugation, followed by isolation of the product and remaining aldehyde by adjusting the pH of the aqueous enzyme reaction mixture to 7.5 and extracting the product into

ethyl acetate.[118] The amine pKa of dopamine is predicted to be 8.93, therefore it should remain in the aqueous layer.[170] Drying of the organic phases by evaporation and resuspension of the resulting residue in 1 M HCl, followed by extraction with dimethyl carbonate (chosen as an environmentally-benign solvent) to remove any remaining aldehyde, resulted in the isolation of the pure product, as determined by NMR analysis, as the HCl salt. The reaction went to a 96% yield (HPLC yield based upon calibration curve of the purified product) and an isolated yield of 88% was achieved, consistent with the smaller scale reactions (92% HPLC yield). ^{1}H-NMR spectroscopic analysis of the product **94** determined that the stereoselectivity of the reaction was mostly retained with a d.r. of 87:13 whereas in the small-scale reaction with purified enzyme, a d.r. of 94:6 was achieved. The reduction in diastereomeric ratio is most likely due to a small amount of background phosphate-mediated reaction arising from the use of cell lysate rather than purified enzyme. It is therefore expected that the d.r. could be improved by using purified enzyme or buffer-exchanged cell lysate instead.

3.2.4 Reactions with enzyme variants

To investigate the effect of NCS variants on the acceptance of α-methyl aldehydes and products formed, several available *Δ*29*Tf*NCS mutants were used as enzyme lysates, namely L76V, A79I, A79F, F80L, M97V, M97L, M97F and Y108F with dopamine **5** and aldehydes **84**, **91**, and **85**. The enzymes were expressed as 50 mL culture volumes by Daniel Méndez-Sánchez using the standard protocol for NCS expression. The enzyme variants all involved single point mutations of active site residues and were designed by a previous PhD student, Benjamin Lichman (Ward and Hailes groups), based upon docking experiments and previous mutagenesis studies to increase the space in the active site for bulkier substrates.[28,22] These enzyme variants were generated by site-directed mutagenesis by B. Lichman and have been shown to exhibit differing activities with different NCS substrates, most notably the mutant A79I has been shown to have superior activities when ketones are used as substrates[36] and L76A has been shown

to affect aldehyde substrate acceptance.[28] All NCS variants were over-expressed (Figure 3.5) and used as clarified cell lysate in enzymatic reactions.

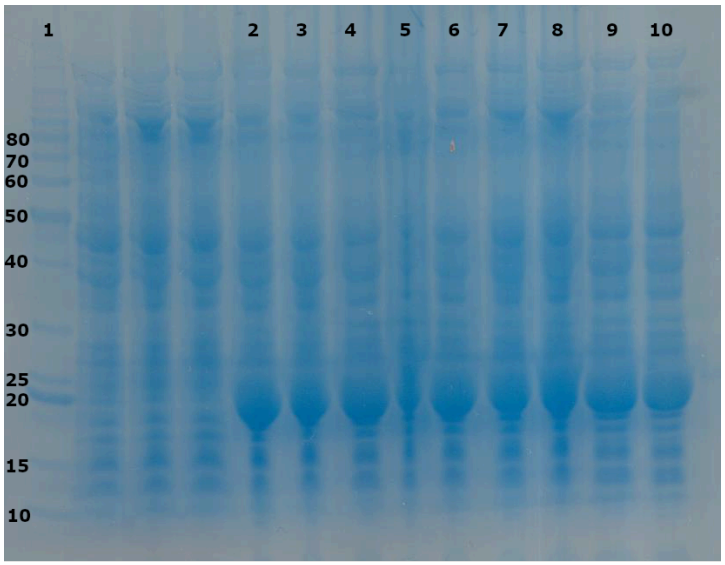


Figure 3.5: SDS-PAGE analysis of Δ29*Tf*NCS single point variant expression. Lanes: 1, Benchmark protein ladder (masses given in kDa). 2. Wild type. 3. L76V. 4. A79I. 5. A79F. 6. F80L. 7. M97V. 8. M97L. 9. ΔM97F. 10. Y108F.

First, WT-*Δ*29*Tf*NCS was used in the reaction and no difference in reactivities or the products formed was found compared to with *Δ*33*Tf*NCS. For all three aldehydes, the mutant A79I showed higher product yields by HPLC compared to WT and the other NCS variants (Figure 3.6), a characteristic that has been observed with ketones to an even greater extent.[36] This mutation has been rationalised to increase conversions of hydrophobic substrates due to enhanced hydrophobicity in the active site resulting in increased substrate affinity. For the aldehydes **84** and **85**, M97V showed increased yields compared to the wild-type enzyme, with complete conversions observed with both. M97 mutants have been used with ketone substrates where decreased yields were noted with a reduction in steric bulk at this position i.e. poorer conversions with M97V compared with the wild-type.[36] Interestingly, the opposite trend was observed here with aldehydes **84** and **85** as substrates.

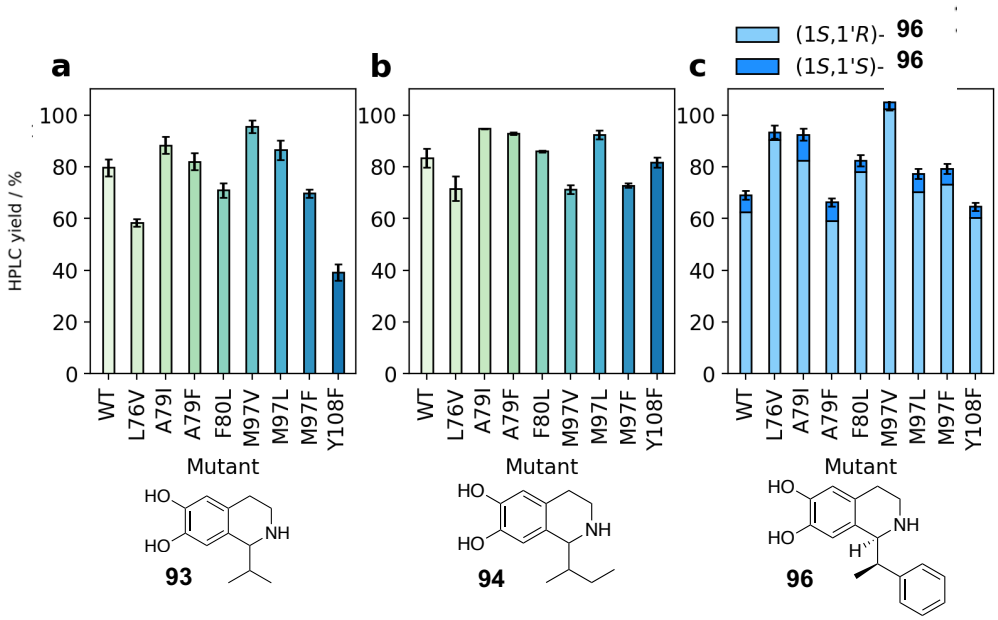


Figure 3.6: **Yields of *Tf*NCS reactions with dopamine 5 and aldehydes 84 (a), 91 (b) or 85 (c) to give products 93, 94 and 96, respectively**. *Reaction conditions*: **5** (10 mM), aldehyde (20 mM), sodium ascorbate (5 mM), Δ29*Tf*NCS lysate of variant (25% *v/v*) in HEPES buffer (100 mM, pH 7.5) with DMSO (10% *v/v*) at 37 ºC for 18 h, 200 μL scale reactions; Yields were determined by monitoring product formation against standards by HPLC (see Appendix). Diastereoisomeric ratios of the major (1*S*,1'*R*) product to the minor (1*S*,1'*S*) product for each reaction product of **96** are; WT, 90:10; L76V, 97:3; A79I, 89:11; A79F, 89:11; F80L, 95:5; M97V, 97:3; M97L 91:9; M97F, 92:8; Y108F, 93:7. Reactions were performed in triplicate and error bars are the standard deviations. Figure adapted from Roddan *et al*.[148]

HPLC analysis of the reactions between dopamine **5** and aldehyde **85** showed that different active site mutants of *Tf*NCS gave different diastereoisomeric ratios of the products. A ratio of 90:10 was observed for wild-type NCS lysate, lower than with purified NCS (94:6), most likely due to the increased background phosphate reaction when using lysate. The two most promising active site mutants observed were L76V and M97V, with diastereomeric ratios of 97:3 observed in the products (**96**) generated. To fully explore the change in reaction selectivity of these two mutants, both mutant enzymes (M97V and L76V) were purified, and reactions performed between dopamine **5** and aldehydes **84**, **91, 85** and phenylacetaldehyde **76**. The three NCS variants (L76V, M97F and M97V) were expressed on a 500 mL scale and purified by Nickel affinity chromatography. SDS-PAGE analysis of the purifications is given in Figure 3.7. The

mutant M97F was also explored to provide a comparison with the M97V mutant (Figure 3.8).

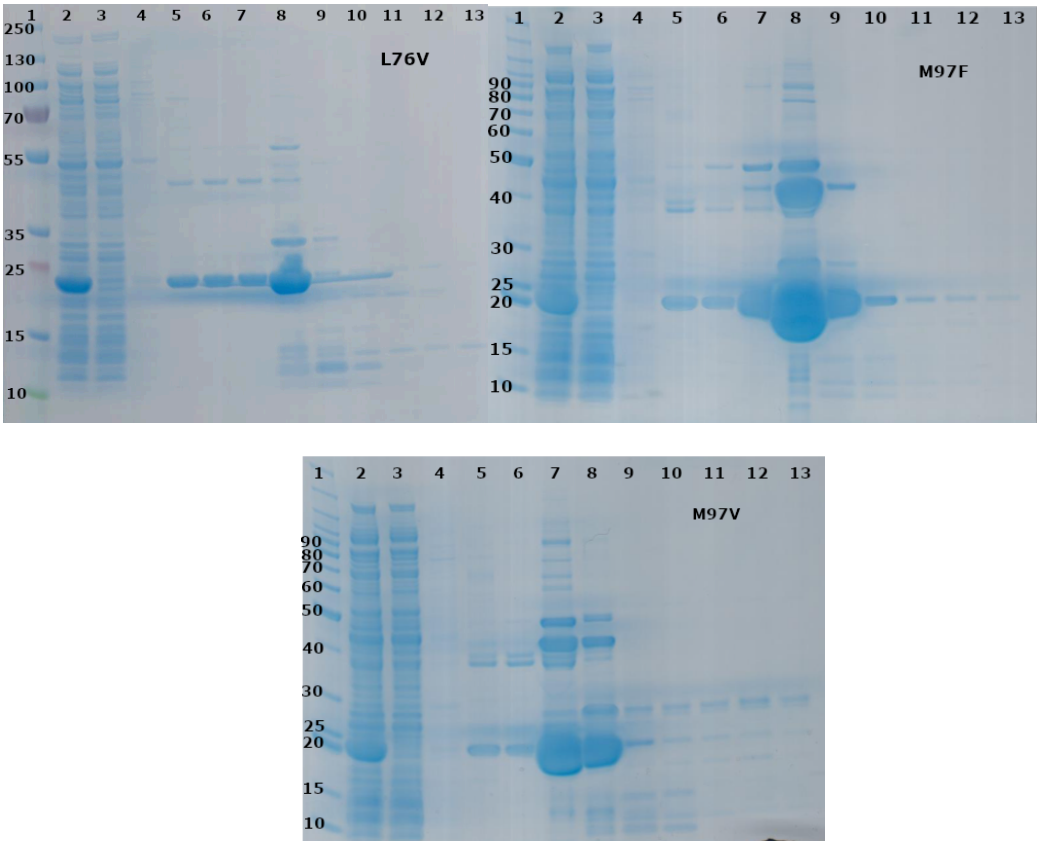


Figure 3.7: SDS-PAGE analysis of His-trap purifications of three Δ29*Tf*NCS single-point variants; L76V, M97F and M97V. Lanes: 1, Benchmark™ Protein Ladder (or Prestained PageRuler Plus™ for L76V), masses given in kDa. 2, clarified cell lysate loaded onto column. 3, flow through with lysis buffer. 4, wash with 6 CV 20 mM imidazole buffer. 5, wash with 6 CV 40 mM imidazole buffer. 6-13, wash with 500 mM imidazole buffer and collected 3 mL fractions. Figure adapted from Roddan *et al*.[148]

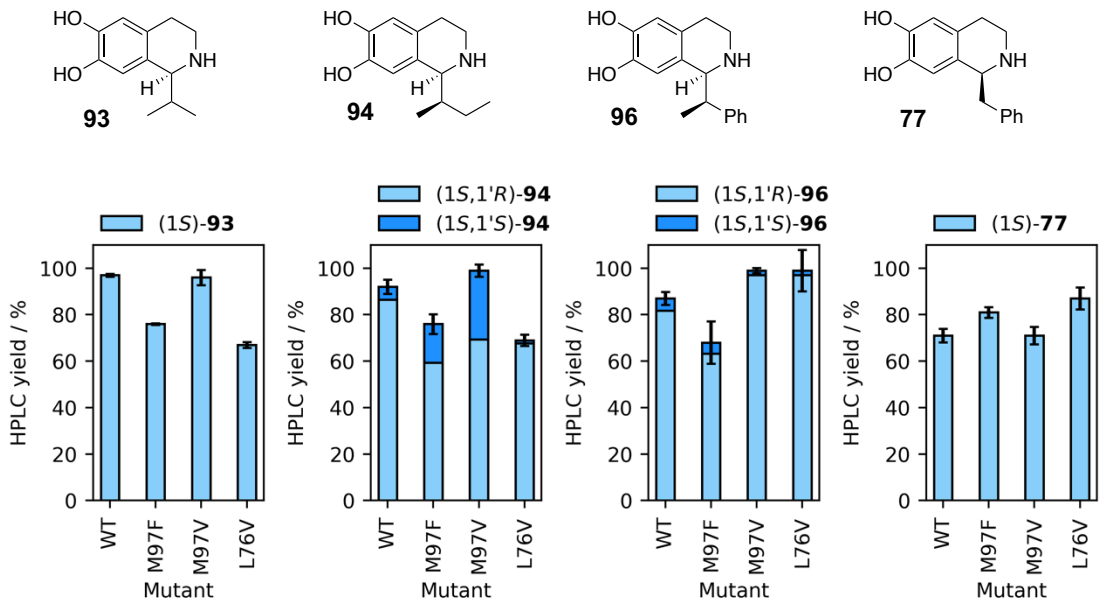


Figure 3.8: **Yields and e.e.s of reactions of products of enzymatic reactions using purified *Tf*NCS active site mutants.** *Reaction conditions:* **5** (10 mM), aldehyde (20 mM), sodium ascorbate (5 mM), purified *Tf*NCS (0.2 mg mL^{-1}) in HEPES buffer (100 mM, pH 7.5) with DMSO (10% *v/v*) at 37 °C for 18 h; Yields were determined by product formation against standards using HPLC (SI Section 9.2). Reactions were performed on a 250 μL scale for **93**, **96** and **77**. Reactions were performed on a 1 mL scale for **94**. e.e.s were determined by chiral analytical HPLC for **93** and **77** (method 3). Diastereoisomeric ratios were determined by ^{1}H-NMR spectroscopy for **94**. The diastereoisomeric ratio was determined by analytical HPLC for **96**. No background reactions were observed, so HPLC yields for a no enzyme control are not shown. Figure adapted from Roddan *et al*.[148]

Aldehyde **76** was used as a substrate with all three mutants to help determine whether the variants were affecting the stereochemistry of the product at the C-1 position. **76** is often used as an alternative to the natural aldehyde substrate, 4-HPAA (**6**), as it is commercially available and less-oxidatively sensitive. As previously reported, the reaction with wild-type *Tf*NCS gave a high e.e. in the product **77**,[22] with only the *S*-enantiomer of product observed by chiral HPLC (Figure 3.9), as did all the NCS variants tested here, implying that there was no effect on the stereochemistry at the C-1 position. The racemic standard of **77** was prepared by a phosphate-mediated Pictet-Spengler reaction between dopamine **5** and phenylacetaldehyde **76** and isolated by preparative-

HPLC (method 8).[31] Comparison with RP-HPLC traces of the same samples determined that for all enzymatic reactions, the peak at 28 – 30 min retention time corresponded to dopamine rather than (*R*)-**77**.

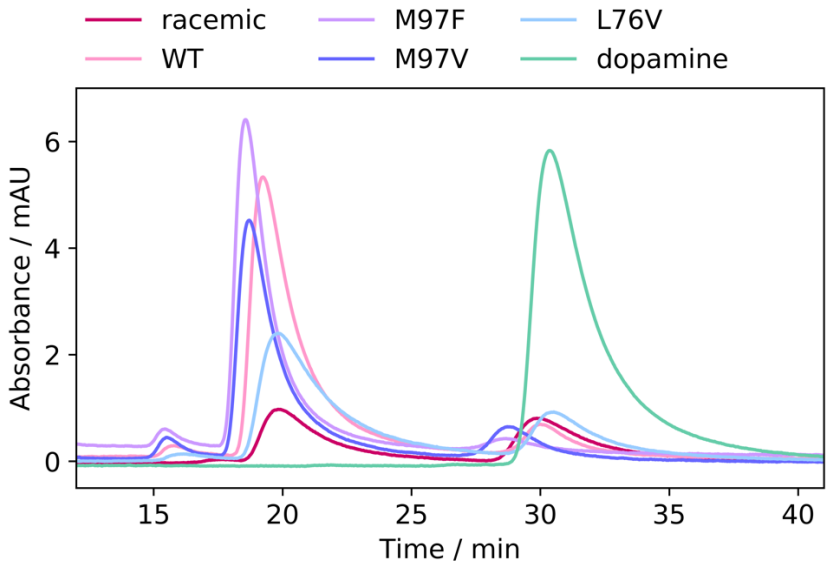


Figure 3.9: **Chiral HPLC analysis of 77 by method 3 using NCS variants as reaction catalysts.** *Reaction conditions for enzymatic reactions:* 10 mM **5**, 20 mM **76**, 5 mM sodium ascorbate, 0.2 mg mL^{-1} purified Δ29*Tf*NCS (wild-type = WT, Δ29*Tf*NCS-M97F = M97F, Δ29*Tf*NCS-M97V = M97V, Δ29*Tf*NCS-L76V = L76V) in HEPES buffer (100 mM, pH 7.5) with 10% *v/v* DMSO at 37 °C for 18 h, 250 μL scale reactions. *Reaction conditions for preparation of racemic* ***77****:* 10 mM dopamine.HCl, 20 mM aldehyde, 5 mM ascorbic acid in phosphate buffer (0.3 M, pH 6) with 50% *v/v* acetonitrile at 60 °C, 10 mL scale. **77** was isolated by preparative HPLC (method 8). A 1 mM solution of dopamine was used and a 5 μL injection volume was used. Figure adapted from Roddan *et al.*[148]

For all reactions with **84** as the aldehyde substrate, a high enantiomeric excess (> 95%) in the product **93** was also retained. The enantiopurity of **93** also demonstrated that the variants were not affecting the stereochemistry at the C-1 position of the products with α-methyl aldehydes as substrates. Yields with M97F and L76V and **84** as the aldehyde substrate were however slightly decreased compared to the wild type (Figure 3.10).

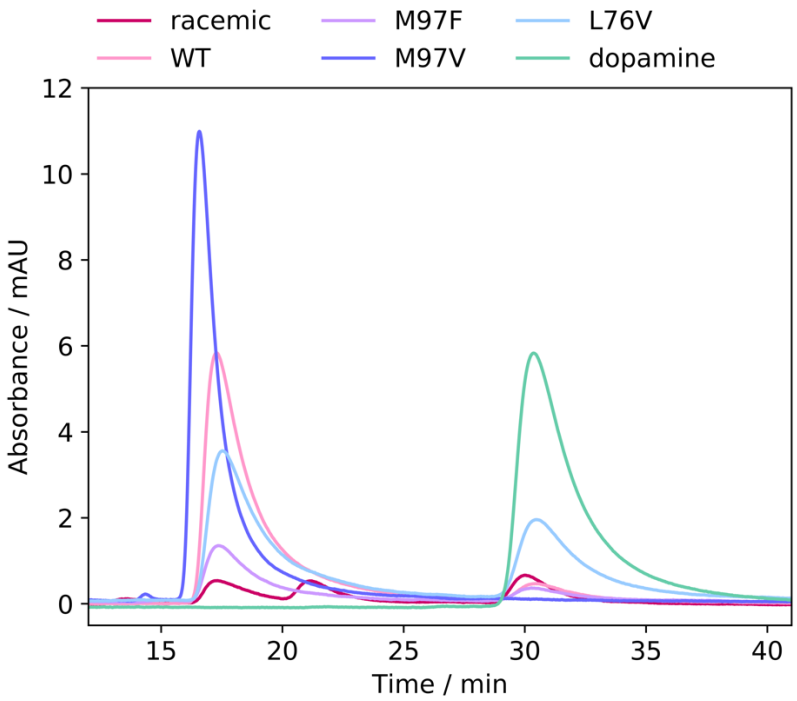


Figure 3.10: Chiral HPLC analysis of 93 using HPLC method 3) using NCS variants as reaction catalysts. *Reaction conditions for enzymatic reactions:* 10 mM **5**, 20 mM **84**, 5 mM sodium ascorbate, 0.2 mg mL^{-1} purified Δ29*Tf*NCS (wild-type = WT, Δ29*Tf*NCS-M97F = M97F, Δ29*Tf*NCS-M97V = M97V, Δ29*Tf*NCS-L76V = L76V) in HEPES buffer (100 mM, pH 7.5) with 10% *v/v* DMSO at 37 °C for 18 h, 250 μL scale reactions. *Reaction conditions for preparation of racemic **93**:* 10 mM **5**, 20 mM **84**, 10 mM ascorbic acid in phosphate buffer (0.3 M, pH 6) with 50% *v/v* acetonitrile at 60 °C, 10 mL scale. A 1 mM solution of dopamine was used with an injection volume of 5 μL. The peak at 21 min corresponds to a small amount of *ortho*-THIQ formed. Figure adapted from Roddan *et al.*[148]

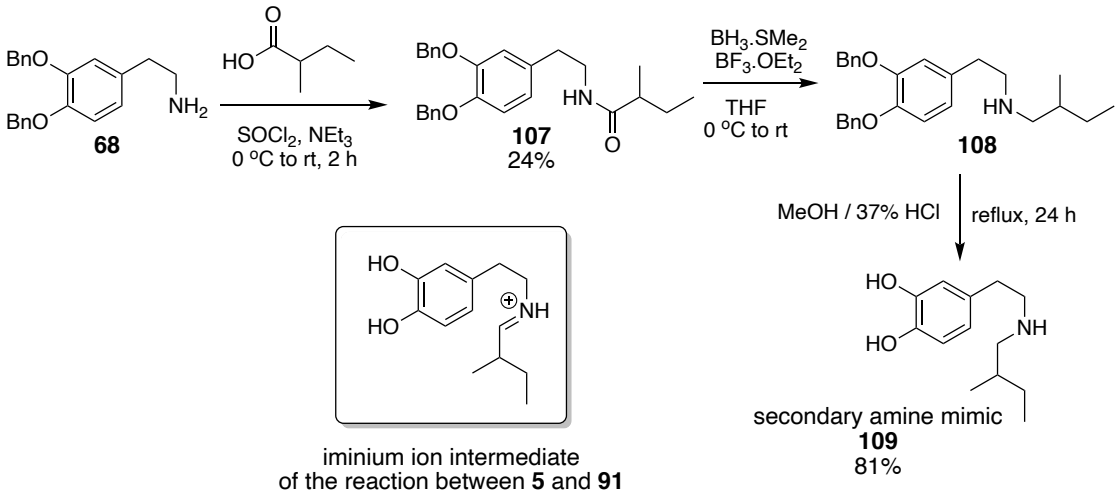
Most notably, for all 3 α-methyl aldehydes tested as substrates, the products of reactions with the L76V variant showed equal or improved stereoselectivities when compared with the wild-type enzyme. An increased yield was observed with **85** as the aldehyde substrate, potentially as extra space is created in the active site. There was a lowering in d.r. for all products in M97F reactions with the chiral α-methyl aldehydes, suggesting that this mutant is affecting the stereochemistry of the C-1' position. With the bulky, chiral aldehyde **85** however, the diastereomeric ratio in the product **96** was improved compared with the wild type for both M97V and L76V variants, and both gave the product in almost quantitative yields. For both variants, there was also little change in diastereomeric ratio between using purified enzyme compared with enzymatic lysate

(Figure 3.6 vs. Figure 3.8), highlighting that purified enzyme is not essential for gaining a high d.r. in the product for these reactions. Interestingly, when using **91** as the aldehyde substrate, a reduction in the diastereomeric ratios in the products were observed when using *Tf*NCS variants altered at position-97 (M97F and M97V).

3.2.5 Structural studies of NCS reactions with α-methyl aldehydes

Previous structural studies of *Tf*NCS by X-ray crystallography and computational docking have elucidated that a 'dopamine-first' rather than an 'aldehyde-first' is the most probable mechanism of action. The roles of key active site residues were also identified.[28,30] One of the important pieces of evidence gained was from a co-crystallised structure of the wild-type enzyme with a non-productive, secondary amine analogue of the iminium ion intermediate bound in the active site, based on the assumption that this would mimic the behaviour of the iminium ion. The intermediate mimic (**16**) that was designed was based upon the *Tf*NCS reaction between dopamine (**5**) and a methylated analogue of natural aldehyde substrate, 4-HPAA (**6**) to minimise oxidative sensitivity.[30] Therefore, to gain more insight into the acceptance of α-methyl aldehydes by *Tf*NCS and expand the mechanistic understanding of *Tf*NCS beyond the natural substrates, a similar approach was used, investigating the co-crystallisation of α-methyl substituted, secondary amine mimics.

Initially, a reaction mimic was synthesised based on the reaction between dopamine (**5**) and aldehyde **91** because the stereochemical outcome of this reaction with *Tf*NCS had been confirmed experimentally (Scheme 3.6). The mimic was synthesised in an analogous manner to that performed previously (Section 2.2.4).[30] The synthetic route is given in Scheme 3.7.



Scheme 3.7: **Synthetic route towards the secondary amine mimic, 109 of the iminium ion intermediate of the Pictet-Spengler reaction between 5 and 91.**

The synthesis of **68** is further discussed in Chapter 2 and was performed on a 5 g scale so that it could be used for the synthesis of other secondary amine mimics (Sections 2.2.4, 5.2.13, 6.2.4). The amide coupling step was performed using a one-pot method devised by Leggio *et al*., involving the *in situ* activation of the carboxylic acid by conversion to the corresponding acyl chloride by reaction using thionyl chloride.[138] The resulting amide (**107**) was obtained in reasonable yield (24%) and was purified by column chromatography to remove any excess benzyl alcohol from the second reaction step. The reduction of the amide (to give **108**) followed by benzyl deprotection was performed in an analogous manner to that described in the literature, with the final product **109** isolated in 81% yield after purification by preparative HPLC (method 8). **109** was prepared as the TFA salt at 200 mM in DMSO ready for co-crystallisation.

A truncated construct of *Tf*NCS, Δ33*Tf*NCS with residues in the flexible, terminal signal peptides removed (1-33 and 196-210) was used previously and shown to be highly soluble during purification and crystallisable under a range of conditions. A discussion of the expression and purification of Δ33*Tf*NCS is given in Section 2.2.1. Resultant protein crystals can also be formed in a range of conditions and are capable of diffracting to high resolution (Section 2.2.3). During previous crystallographic studies

with the apo structure, the highest resolution structures were found in conditions in the 96-well Structure Screen (Molecular Dimensions). Therefore, initial screens were performed using this commercial screen with the sitting-drop method. The final concentration of **109** used was 10 mM (i.e., 5% v/v) with 12.8 mg mL^{-1} Δ33*Tf*NCS. Crystals were observed under a range of conditions however diffraction was only observed in two conditions, D10 (50 mM KH_2PO_4, 20% v/v PEG 8,000) and F10 (100 mM MES pH 6.5, 12 % w/v PEG 20,000). However, poor resolution was observed (3 – 4 Å). All crystals were cryo-cooled using ethylene glycol (20%) in the mother liquor as the cryoprotectant.

In an attempt to improve the crystallographic resolution, four-corner screens were prepared i.e., optimising the precipitation condition by testing conditions which vary slightly from the original condition. For the D10 condition, a 96-well plate was prepared varying the buffer condition from 40 – 60 mM KH_2PO_4 and 15 – 25% w/v PEG 8,000, and for F10, the buffer condition was varied from 50 – 150 mM MES pH 6.5 and 8 – 16 % w/v PEG 20,000. For the D10 four-corner screen, numerous crystals were obtained and with five different crystals sent to a synchrotron for data collection. Unfortunately, poor resolution was observed for all datasets (2.8 – 3.5 Å) with no positive difference density observed in the enzyme active site. More promising results were observed with crystals from the F10 optimisation screen, with a dataset collected at 2.5 Å in space group $P3_121$. Some positive difference density was observed in the enzyme active site near to K122, where the catechol binds, however the ligand could not be built in convincingly (Figure 3.11) and it is possible that a water network is being observed. This suggested that there were issues with poor ligand occupancy, possibly combined with multiple conformations of the ligand in the active site.

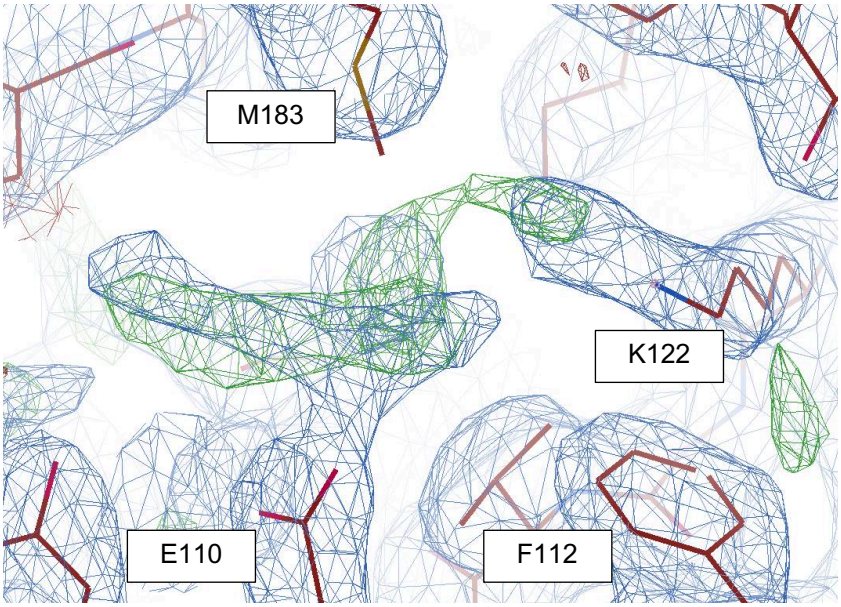
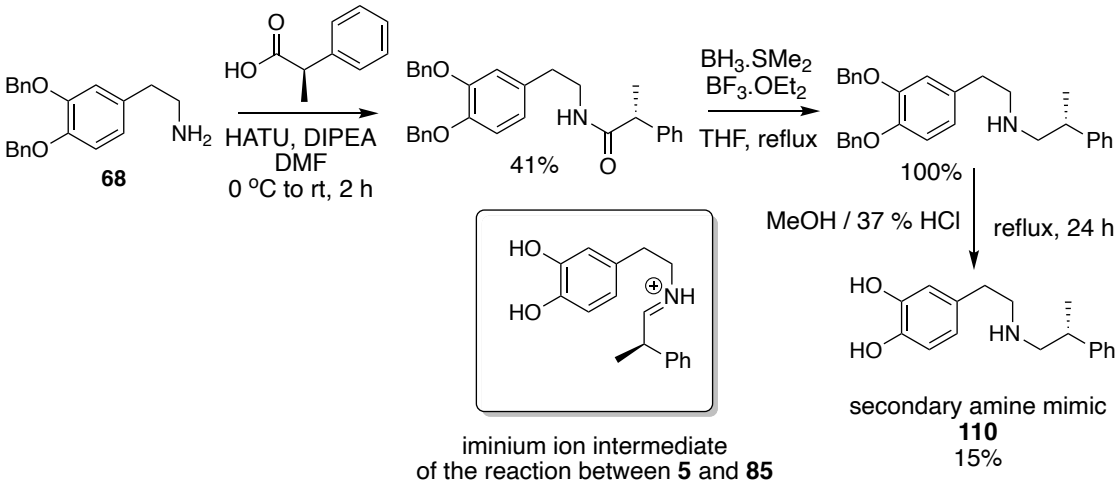


Figure 3.11: **Positive difference density observed in the active site of Δ33*Tf*NCS gained from co-crystallisation with reaction mimic**, **109**. Map is shown at contour level 0.93 rmsd. Structure gained from F10 four-corner screen optimisation. Image generated using COOT.[171] Figure adapted from Roddan *et al.*[148]

The crystals gained for previously co-crystallised structure of Δ33*Tf*NCS with an active site bound reaction intermediate analogue (PDB: 5NON) were obtained via the hanging drop method in 10% w/v PEG 1000 and 10% w/v PEG 8000. Therefore, crystals were prepared using the same method, but instead using the mimic**, 109**. A small amount of electron density was observed in the active site; however, ligand placement was indeterminable. Soaking of the mimic into apo crystals was also unsuccessful, with crystals dissolving upon incubation with the mimic. It was suspected that issues with the binding of this new mimic could be for various reasons. Firstly, a racemic formulation of the mimic was used, therefore, if both enantiomers were capable of binding in the active site, multiple conformations of ligand would be observed, making modelling of the ligand challenging. Also, from enzymatic assays it is known that α-methyl substituted aldehydes are more challenging substrates for NCS, with high substrate loading (10 - 20 mM) required for a successful reaction. The mimic used here, **109**, is based on the reaction between dopamine and an aliphatic aldehyde. The aliphatic chain means that **109** is more flexible than the previously co-crystallised mimic and very high

occupancy and single ligand placement would be required to observe the ligand clearly in the active site. Therefore, with these considerations in mind a new mimic was synthesised, based upon the reaction between dopamine (**5**) and (*R*)-**85**, because this enantiomer of aldehyde was shown previously to be preferentially accepted by the enzyme. It was hoped that having a single enantiomer of mimic and extra electron density from the additional phenyl ring would improve ligand placement. An analogous synthetic route was performed to generate the new mimic (**110**) and is given in Scheme 3.8.



Scheme 3.8: **Synthetic route towards a secondary amine mimic (110) of the iminium ion intermediate of the *Tf*NCS reaction between dopamine 5 and (*R*)-85.**

The synthetic route proceeded as previously discussed except that the amide coupling step was performed under milder conditions, using HATU (Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium) as the coupling reagent, to ensure that no racemisation of the carboxylic acid occurred. The mimic was isolated by preparative HPLC (method 8) and chiral HPLC analysis (method 4) was used to determine whether any racemisation had occurred during the synthesis. The analogous racemic analogue of **110** was therefore also synthesised (*rac*-**110**) to provide a racemic standard, using racemic carboxylic acid rather than the single enantiomer used for the synthesis of (*R*)-**110**. A high e.e. was obtained in (*R*)-**110** (>99%) (Figure 3.12), and (*R*)-**110** was prepared at 200 mM in DMSO, ready for co-crystallisation.

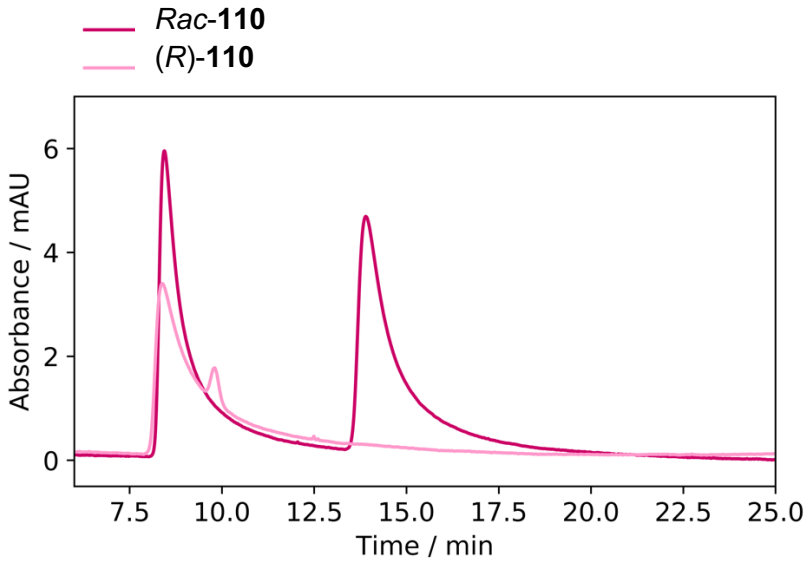


Figure 3.12: **Chiral HPLC analysis of (*R*)-110.** *Rac*-**110** was synthesised in an identical manner. Analysis was performed using chiral HPLC method 4. Figure adapted from Roddan *et al.*[148]

Initial crystallisation trials were performed using four commercially available 96-well screens (Index, PACT, JCSG and Structure). Crystals were observed in a range of conditions and sent to a synchrotron for data collection (Diamond i24) as given in Table 3.2.

Table 3.2: **Conditions in which crystals were observed upon co-crystallisation of mimic (*R*)-110 (10 mM) with Δ33*Tf*NCS (12.6 mg mL^{-1}).** The high-resolution limits of the datasets obtained are also given. MIB buffer: sodium malonate dibasic monohydrate, imidazole, boric acid, PCTP buffer: sodium propionate, sodium cacodylate trihydrate, Bis-Tris propane.

Dataset	***Salt***	***Buffer***	***pH***	***Precipitant***	***Dataset resolution / Å***
1	-	-	-	24% w/v PEG 1500 20% v/v glycerol	2.1
2	0.1 M NH_4OAc	0.1 M Bis-Tris	5.5	17% w/v PEG 10,000	3.3
3	2.0 M $(NH_4)_2SO_4$	0.1 M NaOAc	4.6	-	2.5
4	-	0.1 M MIB	6.0	25% w/v PEG 1500	2.8
5		0.1 M PCTP	6.0	25% w/v PEG 1500	2.9
6		0.1 M Bis-Tris	6.5	28% w/v PEG monomethylether 5,000	2.5
7	-	0.1 M Bis-Tris	6.5	28% w/v PEG monomethylether 2,000	2.4
8		0.1 M HEPES	7.5	25% w/v PEG 3,350	2.6

Although high resolution was achieved for dataset 1, no positive difference density was observed in the enzyme active site, although the putative gatekeeper residue, F112 (discussed further in Section 2.2.6) was present in a different rotameric form to the apo structure. Dataset 2 was not refined as poor resolution was achieved. Dataset 3 resulted in a dataset with space group $P12_11$ with some positive difference density observed in the active site, although the R-factors were high. Positive difference density was also observed in the active site for datasets 4, 5, 6 and 8, with particularly promising results observed with datasets 6 and 8 although poor resolution limited ligand building. No positive difference density was observed with dataset 7.

Therefore, to optimise the crystallisation conditions, a hanging-drop four-corner screen was prepared in a 24-well plate around the conditions used for dataset 6, varying the buffer concentration from 80 mM – 120 mM Bis-Tris pH 6.5 and the precipitation concentration from 18 – 22% w/v PEG-monomethylether 5,000. Crystals only appeared in wells with 22% w/v PEG-monomethylether 5,000 with varying concentrations of Bis-Tris buffer. A range of crystals were sent to a synchrotron (Diamond i24) and one high resolution dataset was collected, at 1.8 Å with the precipitant 22% PEG monomethylether 5,000, 80 mM Bis-Tris pH 6.5. Enzyme, Δ33*Tf*NCS (12.3 mg mL-1 in 20 mM Tris, 50 mM NaCl, pH 7.5) was incubated with (*R*)-**110** (200 mM in DMSO) to give a final mimic concentration of 10 mM. Crystals were grown for 3 days and cryo-protected in the crystallisation condition containing 20% ethylene glycol and 10 mM of the mimic, (*R*)-**110**. A co-crystallised structure of *Δ*33*Tf*NCS with mimic (*R*)-**110** in the active site was gained at 1.81 Å with a single copy in the asymmetric unit (PDB entry: 6RP3, Table 3.3).

Table 3.3: X-ray data collection and refinement statistics for co-crystallised structure of *Δ*33*Tf*NCS with α-methyl substituted reaction mimic, (*R*)-110 bound in the enzyme active site.

	PDB: 6RP3
Space group	$P3_221$
a, b, c (Å)	62.26, 62.26, 73.58
α, β, γ (°)	90.0, 90.0, 120.0
Resolution range (Å)	43.49-1.81 (1.84-1.81)
Total number of observation	159646 (4158)
Total number unique	15507 (745)
Completeness	99.9 (98.4)
Multiplicity	10.3 (5.6)
<I/σ(I)>	10.1 (0.98)
$CC_{1/2}$	0.99 (0.64)
R_{merge}	0.090 (1.949)
R_{pim}	0.029 (0.842)
Molecule per ASU	1
Refinement	
Resolution Range (Å)	43.49-1.81 (1.856-1.809)
R_{work}	0.203 (0.315)
Rf_{ree}	0.245 (0.358)
Reflection, working	14703
Reflection, free	763
Average B factor	28.8
Rmsd bond angle	1.816
Rmsd bond length (Å)	0.014
Ramachandran plot	
Preferred region (%)	95.5
Allowed region (%)	3.2
Outliers (%)	1.3

Data processing was performed using the CCP4i2 package. Images collected were integrated using XDS Dials.[172] Phases were determined by molecular replacement, using Phaser[173] using an apo dataset of the same NCS construct (Δ33*Tf*NCS, wild-type) in the same space group ($P3_221$) as the search model. COOT[171] was used for

model building of the protein and the ligand and REFMAC5[174] used for refinement with TLS refinement used to account for anisotropy in the dataset. Groups of atoms were assumed to act as a rigid body, of which there is a mean squared displacement, therefore modelling the anisotropic motions of atoms and thus can lead to improvements in observed R-factors in structure refinement. The final Rwork/Rfree factors after refinement were 0.20/0.25 with three molecules of ethylene glycol (arising from the cryo-protectant used) and fourty-four water molecules built in.

Compared to the previous apo structure of *Δ*33*Tf*NCS (PDB: 5N8Q), the overall fold of NCS was retained (RMSD C_α= 0.587). The only notable differences in active side residue conformations were with E110 and M97. The mimic density in the active site was less convincing than with the previously co-crystallised structure (PDB: 5NON) suggesting partial occupancy and other, minor multiple conformations of the ligand. A Polder[175] map of the ligand density showed similar density features, with some uncertainty in the placement of the dopamine ring (Figure 3.13). However, the 'aldehyde-end' of the mimic could be observed relatively clearly and so this was used to help understand why these substrates are accepted and why activity varies with the single-point mutants.

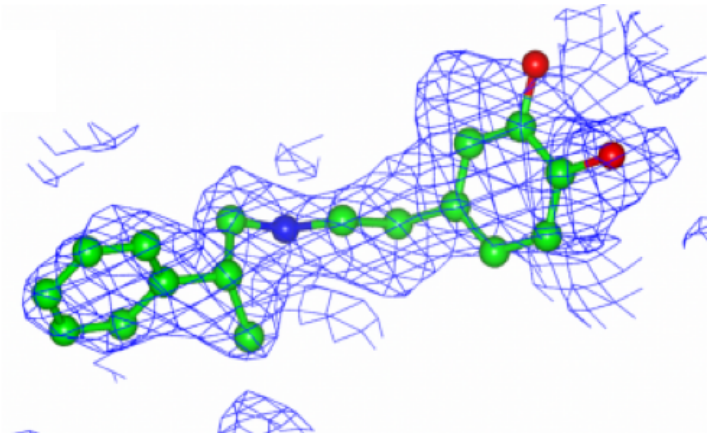


Figure 3.13: **Electron density of (*R*)-110 in the active site at contour level 0.81 RMSD.** Figure adapted from Roddan *et al*.[148]

In this structure, the ligand is placed with both hydroxyl groups of the catechol hydrogen bonded to Lys122 (with distances of 3.28 and 3.02 Å respectively) in a linear configuration, likely reflecting a pre-cyclisation conformation (Figure 3.14). Previous studies involving the active site mutagenesis of key residues in NCS have shown that K122 is essential for catalytic activity and is required for the deprotonation of the *meta*-hydroxyl.[7,28] The hydrogen bonding distances are further away than in PDB:5NON and the mimic appears to be further buried into the active site. The interaction with both hydroxyl groups is interesting considering that lower conversions were observed when dopamine (**5**) was exchanged for *meta*-tyramine (**51**) as the amine substrate (Table 3.1). This suggests that this observed additional hydrogen bonding interaction of the *para*-hydroxyl to K122 may be required for holding the amine substrate in the active site leading to increased likelihood of a successful reaction.

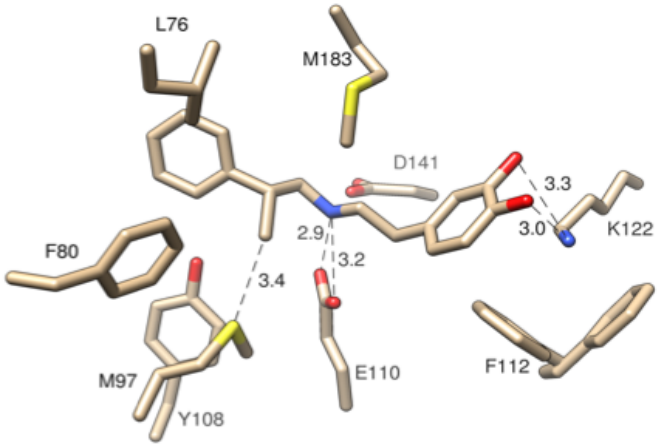


Figure 3.14: **Location of mimic (*R*)-110 in the active site**. Dotted lines show key distances or hydrogen-bonding interactions. Associated numbers are distances given in Å. Figure adapted from Roddan *et al.*[148]

A hydrogen bonding interaction can also be observed between the secondary amine of the mimic and Glu110, which is in a rare conformation, observed in only 1% of glutamic acid residues in the PDB, whereas in the previous structure, the amine was bound to both E110 and D141. This interaction is consistent with the 'dopamine-first' mechanism (detailed in Section 1.3.4) and we are likely observing another 'snapshot' of the mechanistic process. F112 is present in two different conformations in the novel

structure (defined at 1:1 occupancy), as in the first holo structure of *Tf*NCS (PDB: 2VQ5)[27]. In subunit B of 2VQ5, the phenylalanine conformation is uncommon, being present in only 1% of phenylalanine residues in the PDB. In structure 5NON, Phe112 is predominately in the conformation predicted by MD simulations and in subunit A of 2VQ5.[28] It is not known whether this unusual phenylalanine conformation is relevant to the NCS mechanism, perhaps suggesting an inhibitory mode with the residue acting as a 'gatekeeper', or if it is simply a crystallographic artefact.

The density of the 'dopamine-end' of the mimic was less clear so the ligand was built in with both hydroxyl groups hydrogen-bonded to K122, reflecting a likely mixture between the two potential conformations, productive and inhibitory. The phenyl group of the 'aldehyde-end' of the mimic sits close to L76. Mutation of this residue to valine resulted in improved stereoselectivity for NCS reactions with α-methyl aldehydes. The 'aldehyde-end' of the mimic can be observed clearly and the methyl group of the mimic appears to sit close to M97, forcing the residue into a previously unseen conformation for this residue in NCS structures. Reducing bulk at this position by mutation to a valine residue, lowers d.r.s for the reaction with aldehyde **91**, thus providing extra space for an ethyl group and so both enantiomers of the aldehyde are more readily able to react with dopamine in the active site. However, this does not provide enough extra space for the phenyl group of **85** and so one enantiomer is accepted preferentially. An increased d.r. in the product **96** compared to the wild type is likely due to improved binding of the aldehyde from this extra space provided and resulted in increased reaction yields. Conversely, the mutant M97F reduced yields and d.r.s in all products, likely due to lack of space in this region of the active site. The mutant L76V improved d.r.s with both chiral aldehydes, possibly due to altering the shape of the active site cavity so that binding of the (*R*)-enantiomer is preferential.

To further confirm the preferential acceptance of the (*R*)-enantiomer of α-methyl substituted aldehydes, molecular dynamics simulations have been performed by Gudrun Gygli.[148] This work investigated the binding of dopamine **5** and the (*R*)- and (*S*)-enantiomers of 2-phenylpropanal **85**, into the *Tf*NCS active site, analysing near attack complexes (i.e. measuring relevant distances between the two substrates, which when below a 4.0 Å, suggest that the substrates are orientated to lead to a productive reaction). The simulations concluded that the (*R*)-enantiomer of aldehyde **85** is preferred to the (*S*), with a ratio of 95:5, closely matching experimentally observed d.r.s with these substrates (94:6, Scheme 3.3). The results were also consistent with the 'dopamine-first' mechanism of NCS whereby dopamine first binds to the active site, followed by the aldehyde.[28,30] These simulations have therefore provided supporting data that a kinetic resolution of the α-methyl-substituted aldehyde was occurring and that the (*R*)-enantiomer of aldehyde was accepted preferentially.[148]

3.3 Conclusions

In summary, this work shows that despite previous reports, α-methyl-substituted aldehydes are indeed accepted as substrates by *Tf*NCS using both dopamine and a variety of analogues as the amine substrate. Products were generated in high yields (up to 97%) with high e.e. (>99%), or d.r.s (up to 94:6) with chiral α-methyl-substituted aldehydes, providing a sustainable, facile route to these compounds. For chiral α-methyl-substituted aldehydes, the THIQ products were generated with two defined chiral centres, with (1*S*,1′*R*)-stereochemistry as determined by the synthesis of two single enantiomeric aldehydes. Active site mutants could improve the d.r.s in the products (up to 98:2) with chiral α-methyl-substituted aldehydes and improved the reaction yields with several aldehydes. The enzymatic reaction was confirmed to occur via a mechanism that involves dopamine binding first to the active site, from a co-crystallised structure of Δ33*Tf*NCS with a reaction intermediate mimic in the active site. Computational studies (performed by Gudrun Gygli) further confirmed the occurrence

of the kinetic resolution of racemic α-methyl-substituted aldehydes by *Tf*NCS, whereby the (*R*)-enantiomer of aldehyde is accepted preferentially.[148]

In future work, confirmation of the stereochemistry at the C-1 and C-1' positions of the THIQ products generated could be gained from small-molecule X-ray crystallography of the isolated products from a *Tf*NCS-catalysed reaction between dopamine and a racemic α-methyl substituted aldehyde. Two enzyme variants, L76V and M97V, were particularly promising for improving d.r.s and conversions in the products. Therefore, further studies could involve mutagenesis at these two positions, exploring increasing and decreasing the steric bulk of the side chains, or testing these mutants with other α-substituted aldehydes as substrates.

Chapter 4: Chemoenzymatic Cascades towards Methylated Tetrahydroprotoberberine and Protoberberine Alkaloids

4.1 Introduction

4.1.1 Accessing natural product drugs

For millennia, humans have been exploiting the medicinal and recreational benefits of the many bioactive molecules that are synthesised by plants and nearly half of all licensed drugs are based upon natural products.[176] However, isolating natural products from plants is difficult; often the desired compound is produced in very low quantities, alongside other structurally similar metabolites. This means that challenging, multi-step purification procedures are required, often resulting in poor product yields.[177] Total synthesis routes to natural products have been successful in some cases and can be commercially viable.[178] However, for some products with high molecular complexity, chemical synthetic routes are often unattainable or not cost effective.

Advances in synthetic biology have led to the production of a variety of natural product drugs *in vivo* via fermentation processes but achieving high enantiopurity in the products has proved challenging with some classes of compounds.[122,127] Another approach is the use of the relevant recombinantly expressed enzymes to mimic the biosynthesis *in vitro*. The use of such over-expressed enzymes provides regioselective and stereoselective control of the reactions involved, can minimise side-reactions and simplify the purification of the desired product. Enzymes are also generally more environmentally friendly than traditional organic synthetic routes, with reactions performed under mild conditions (often aqueous, with low temperatures and pressures) and thus are being increasingly employed in industrial chemical syntheses as the pharmaceutical industry aims to transition to a lower-carbon economy.[179]

Furthermore, the use of both biocatalytic enzymes and traditional synthetic chemistry in cascade processes can lead to the generation of highly complex products in fewer

steps and with more facile purification processes than a solely *in vivo* or for an organic synthetic route. These strategies have led to the production of a range of natural and non-natural products *in vitro* to produce naturally occurring alkaloids and relevant analogues in significant yields.[120,128,180]

4.1.2 Unique alkaloids isolated from the *Corydalis* genus

Flowering plants in the *Corydalis* genus are native to high-altitude grasslands in China. Protoberberine alkaloids isolated from *Corydalis* plants are unique amongst isolated protoberberine alkaloids as they possess a methyl group at C-13. A range of 13-methyltetrahydroprotoberberine (13-MeTHPB) and 13-methyl-protoberberbine (13-MePB) alkaloids (**90, 111-114**) isolated from *Corydalis* plants are given in Figure 4.1.

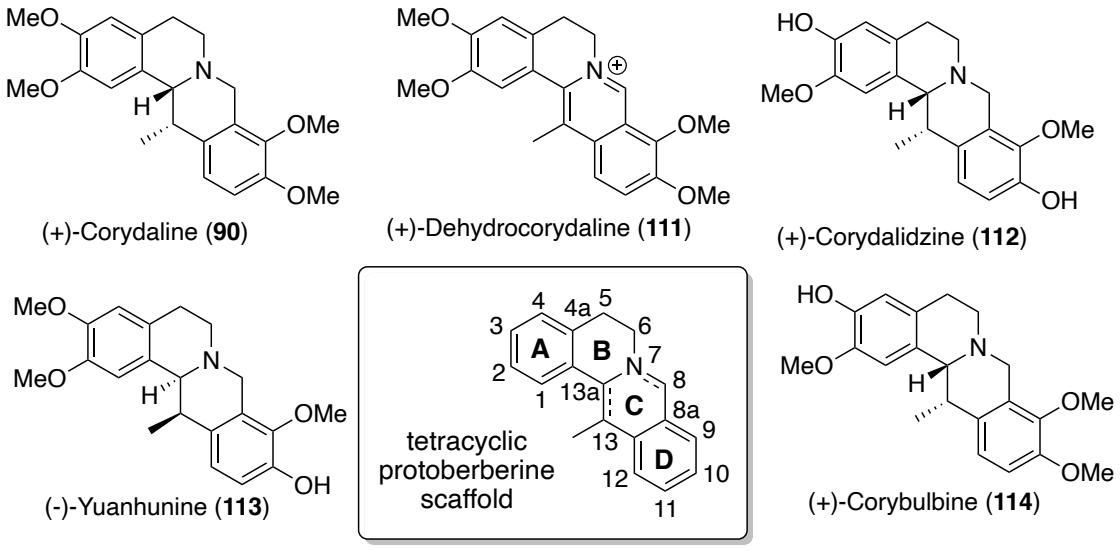


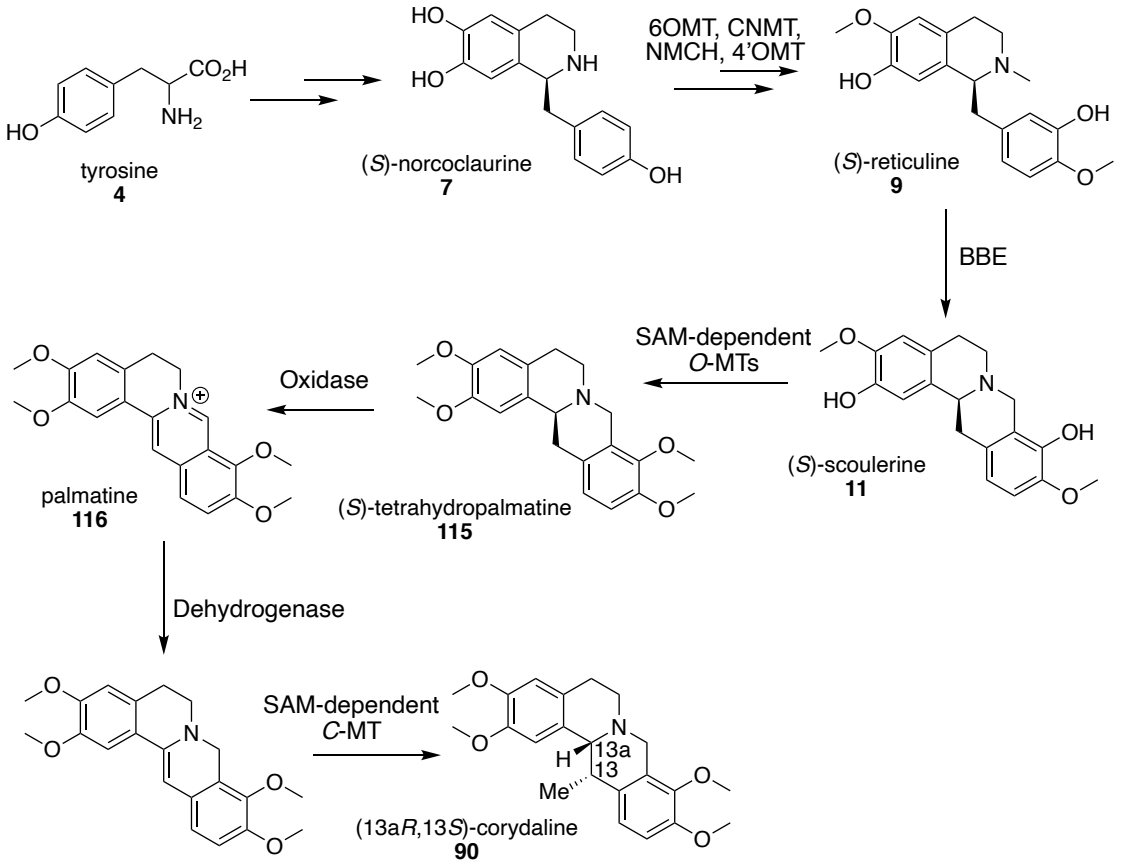
Figure 4.1: Examples of 13-methyl-tetrahydroprotoberine and 13-methyl-protoberberine alkaloids isolated from plants of the *Corydalis* genus.

Many of these compounds have been shown to have promising biological activities. Corydaline (**90**) has been shown to act as an inhibitor of reverse transcriptase activity[181] and inhibit the replication of enterovirus 71, one of the major causative agents of hand, foot and mouth disease.[182] Dehydrocorydaline (**111**) can act as a dopamine D-1 receptor antagonist, so is a promising target for neuropsychiatric disorders[183] and a

range of other 13-MePB alkaloids have been shown to have anti-hepatitis B activities.[184] The extracts of *C. yanhusuo* are used in Chinese traditional medicine for cardiac arrythmia, rheumatism and gastric ulcers.[183] The ethanolic extracts of *C. tuber* are also licensed in South Korea for the treatment of functional dyspepsia.[185]

4.1.3 Biosynthesis of 13-MeTHPB alkaloids

The biosynthesis of 13-MeTHPB alkaloids was investigated in the 1960's and 1970's, aided by the crystal structure of corydaline (**90**) and thus the determination of stereochemistry at both chiral centres.[186] The hypothesised biosynthetic pathway towards corydaline (**91**) is given in Scheme 4.1. Corydaline has been determined to have (*R*) stereochemistry at the C-13 position, so it was initially assumed that the stereochemistry at this position would arise from (*R*)-reticuline (**21**).[187,188] However, this was disproved by MacLean *et al*. who, using feeding studies, determined that palmatine (**116**) was in fact the immediate precursor to corydaline (**90**), rather than (*R,S*)-tetrahydropalmatine (**115**). This was also confirmed by feeding studies performed by Holland *et al*., where it was shown that the C-13 methyl group is incorporated from methionine after the protoberberine scaffold is formed.[189] The biosynthesis of other 13-MeTHPBs (such as those given in Figure 4.1) is unknown, however it is likely that a similar pathway is involved.



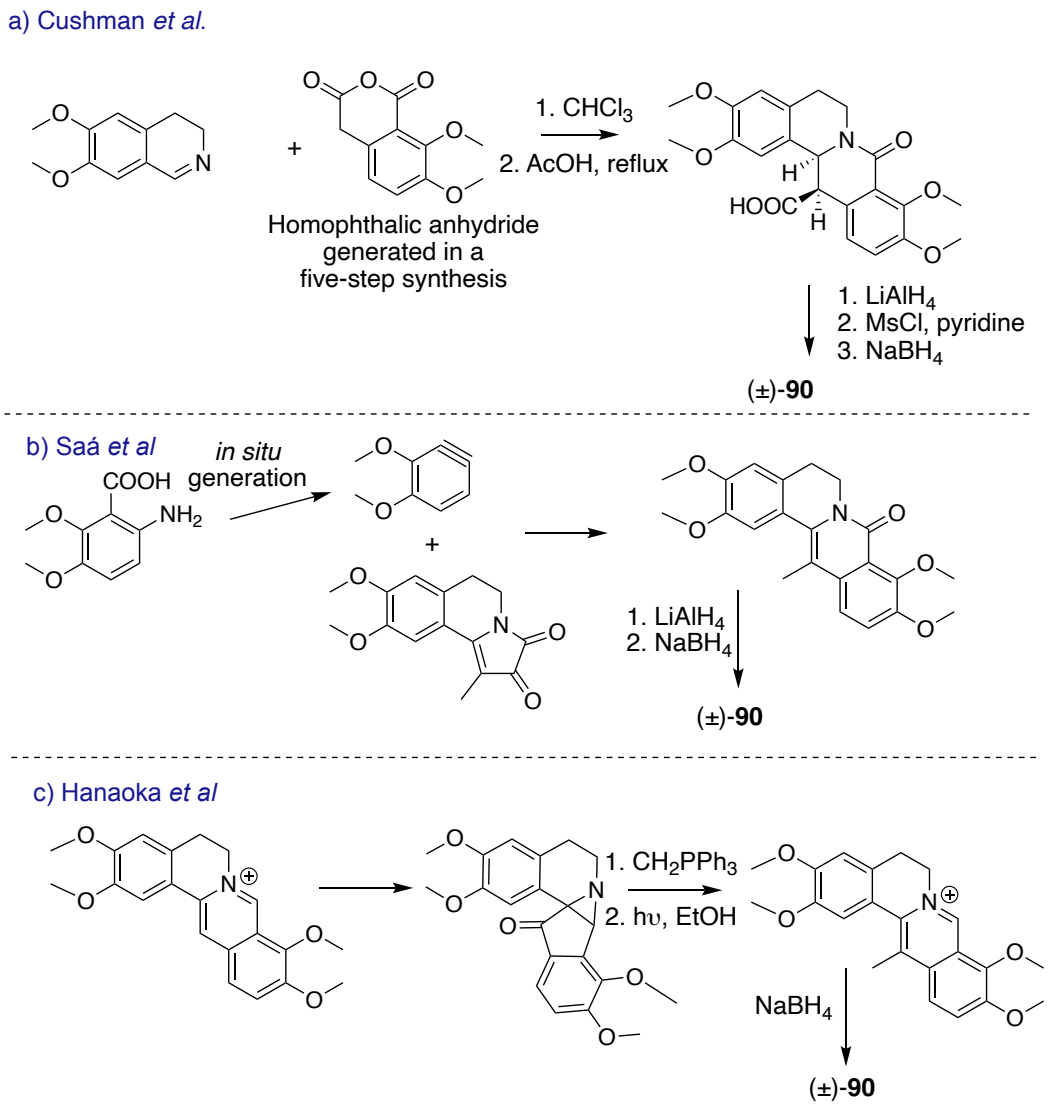
Scheme 4.1: The hypothesised biosynthetic pathway to corydaline and related alkaloids.

The first committed step in the biosynthesis of benzylisoquinoline alkaloids (BIAs) is the formation of (*S*)-norcoclaurine (**7**) by the Pictet-Spenglerase enzyme, norcoclaurine synthase (NCS). This reaction forms the main benzylisoquinoline scaffold and norcoclaurine synthases have been identified in *Corydalis* plants.[20] The NCS in *C. saxicola* has been recombinantely expressed in *E. coli* and shown to generate (*S*)-norcoclaurine in high conversion (90%) and e.e. (>98%).[190] Subsequent methylation steps, performed by regioselective *O*- and *N*- methyltransferases (6OMT, CNMT, NMCH, 4'OMT) give (*S*)-reticuline (**9**).[1] The flavin-dependent, berberine bridge enzyme (BBE) then forms a C-C bond between the *N*-methyl and the aryl ring, giving (*S*)-scoulerine (**11**). The enzyme provides regioselectivity in the ring closure step.[74] Subsequent methylation steps by SAM (*S*-adenosyl L-methionine) dependent methyltransferase enzymes, followed by oxidation of the C-ring gives palmatine (**116**). Reduction with a dehydrogenase followed by stereoselective methylation with a regio-

and stereoselective, SAM-dependent *C*-methyltransferase gives corydaline (**90**), with the stereocentres defined as (13a*R*,13*S*). This is hypothesised to occur in a single step, catalysed by the putative enzyme, corydaline synthase.[191]

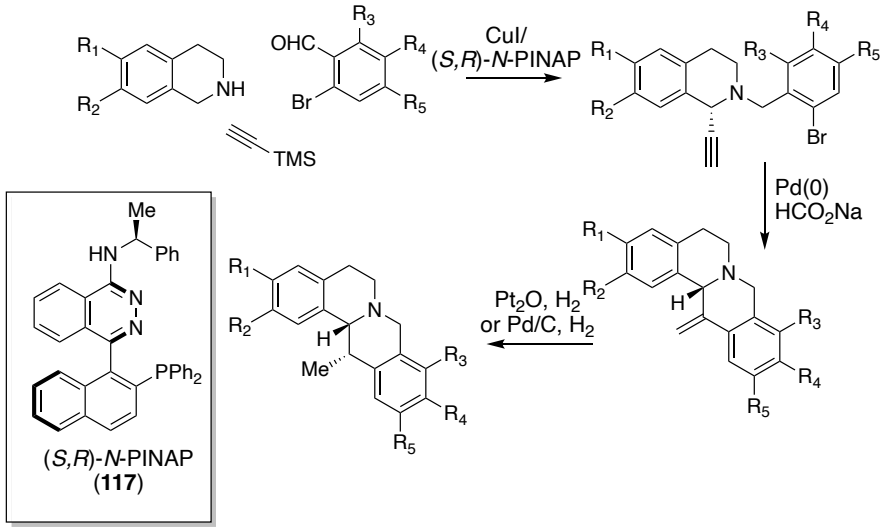
4.1.4 Chemical routes towards 13-MeTHPB alkaloids

To avoid the challenges associated with isolating alkaloids from natural sources, chemical strategies have been devised to synthesise 13-MeTHPBs. The first total synthesis of (±)-corydaline (**90**) was performed by Späth *et al*.[192] starting from papaverine (**10**), however complicated isolation procedures were required and a very low yield of product (0.04%) was isolated. An improved synthesis was later performed by Cushman *et al*. in 1977, with the key synthetic step to form the tetracyclic scaffold involving the reaction between a homophthalic anhydride with a 3,4-dihydroisoquinoline (Scheme 4.2a).[193] This forms the tetracyclic protoberberine scaffold, and *syn* geometry of the two hydrogens at C-13 and C-13a in the product is achieved by performing the reaction under reflux in acetic acid. This is followed by reduction of the amide and carboxylic acid groups, mesyl protection of the alcohol and borohydride reduction to cleave the C-O bond, thus yielding (±)-corydaline (**90**). An alternative route to the same compound has also been devised by Saá *et al*. involving a benzyne cycloaddition reaction (Scheme 4.2b).[194] Other routes have been developed from the corresponding protoberberine alkaloids involving a Wittig reaction followed by a photochemical electrocyclic reaction (Scheme 4.2c).



Scheme 4.2: Synthetic routes developed towards racemic 13-MeTHPB alkaloids.

The first asymmetric total synthesis towards these compounds was developed by Zhou *et al*. in 2017, involving a CuI-catalysed reaction between an alkyne, a 2-bromobenzaldehyde derivative and a THIQ. (*S,R*)-PINAP (**117**) is used as the ligand which was synthesised in four-steps.[195] This was followed by a palladium-catalysed reductive carbocyclization and an asymmetric hydrogenation step (Scheme 4.3).[196] In three steps, this route generated 12 naturally occurring 13-MeTHPBs in good yields (47-64%) and e.e.s (91-96%).



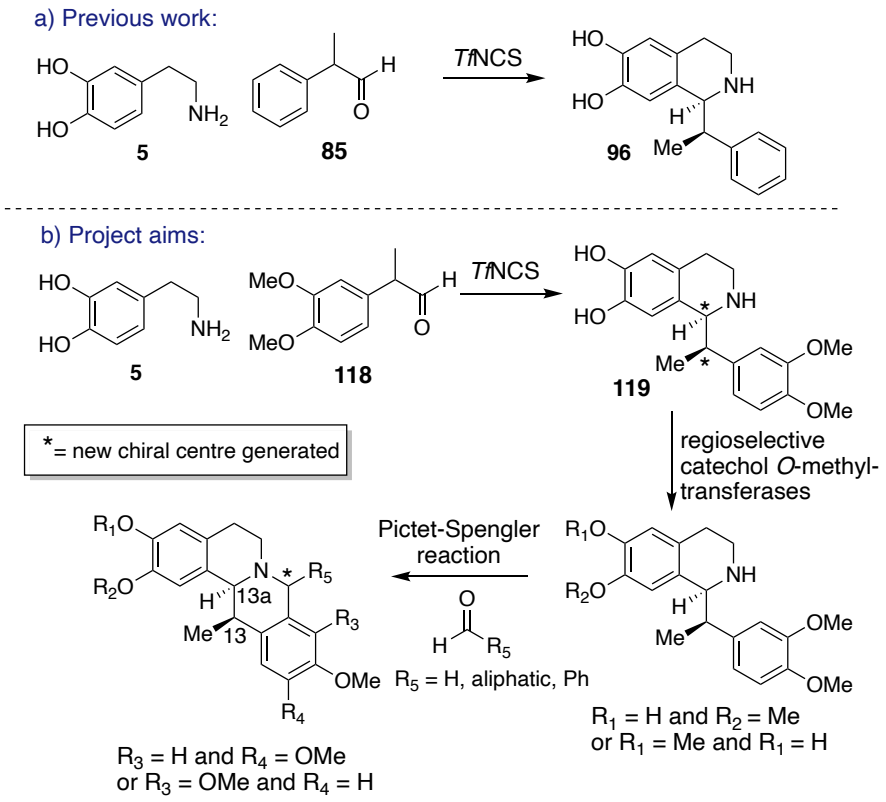
Scheme 4.3: A three-step route towards 13-MeTHPB alkaloids. Route developed by Zhou *et al*.[196]

4.1.5 Biocatalytic approach

Previous work, described in Chapter 3, demonstrated that the Pictet-Spenglerase, norcoclaurine synthase (NCS) can accept a variety of α-methyl substituted aldehydes as substrates, leading to a kinetic resolution of the aldehyde if a racemic aldehyde is used. This led to the production of a range of (1*S*,1'*R*)-THIQs in high diastereomeric ratios (up to 98:2) and yields (up to 96%), using a single-point variant of NCS, *Tf*NCS-M97V (Scheme 3.3). It was hoped that using this enzyme catalysed step, followed by other biocatalytic or chemical steps to generate a range of analogues of the 13-MeTHPBs isolated from *Corydalis* plants. The products generated would have opposing stereochemistry to those isolated from plants, thus providing routes to useful analogues of the isolated natural products for drug discovery purposes.

The aldehyde, 2-phenylpropionaldehyde (**86**), was shown to be well-tolerated as a substrate by *Tf*NCS[148] (Scheme 4.4a) so it was hoped that a non-oxidatively sensitive, methoxy-substituted analogue, 2-(3,4-dimethoxyphenyl)propanal (**118**) could instead be used as the NCS substrate (Scheme 4.4b). This step would provide stereoselective control to generate the two chiral centres in the product, **119** and the subsequent

alkaloids generated. This could then be followed by regioselective methylation reactions, using catechol *O*-methyltransferases to functionalise the catechol towards corydaline-like, 13-MeTHPBs. A range of catechol *O*-methyltransferases have been identified in the literature and work in our group has identified that some benzylic THIQs are accepted as substrates, providing the opportunity to extend the substrate scope towards 1'-Me-THIQs.[197] The THIQ scaffold generated can then undergo another Pictet-Spengler reaction with an aldehyde, to give the desired tetracyclic 13-MeTHPB scaffold (Scheme 4.4b). Analogous chemical Pictet-Spengler reactions have already been employed with simple benzylic THIQs as substrates to generate THPB alkaloids so it was hoped that the substrate scope could be extended here.[117,198] The synthetic route proposed for this work is given in Scheme 4.4b.

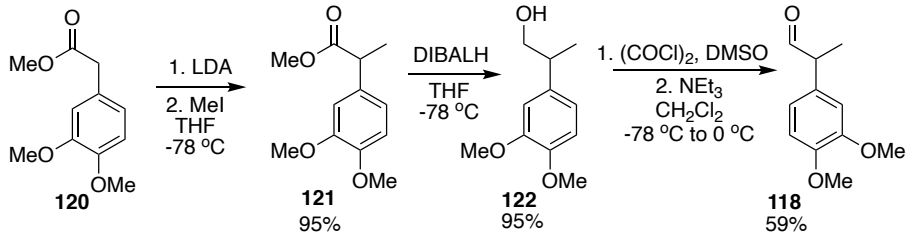


Scheme 4.4: **The acceptance of α-methyl substituted aldehydes by NCS may provide a route towards 13-MeTHPBs a**. Previous work demonstrated that *Tf*NCS-catalysed reactions could be used to generate (1*S*,1'*R*)-THIQs in high yields and diastereomeric excesses. **b**. The proposed project aims, to generate a range of novel (13*R*,13a*S*)-MeTHPBs in three-steps.

4.2 Results and Discussion

4.2.1 Synthesis of 2-(3,4-dimethoxyphenyl)propanal (118)

The 3,4-dimethoxylated analogue of 2-phenylpropanal (**85**), was synthesised in three-steps, starting from the corresponding ester without the α-methyl group. The synthetic route used was based upon a literature procedure and is given in Scheme 4.5.[199]



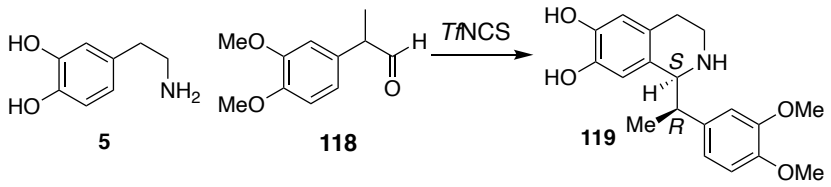
Scheme 4.5: Synthesis of the α-methyl substituted aldehyde, 2-(3,4-dimethoxyphenyl) propanal (118).

The first step, the α-methylation of **120**, was performed on a 1 g scale. The ester was deprotonated using lithium diisopropylamide (LDA), to give the enolate. Iodomethane was then added to provide an electrophilic methyl group and give **121**. Two equivalents of iodomethane were used to ensure that the reaction went to completion and di-methylation was unlikely as after mono-methylation at the α-position, electron donation from the methyl group via the inductive effect makes the enolate less nucleophilic. The reaction proceeded to completion in 1.5 h and the product, **121** was isolated by silica column chromatography. Attempts to perform this reaction on a larger scale (2 g of **120**) were not as effective; a mixture of **120** and **121** were obtained and were challenging to separate via silica chromatography as the two compounds have similar R_f values. Reduction of the ester, directly to the aldehyde using DIBALH has been reported using one equivalent of reducing agent, so reactions were initially attempted using these conditions.[200] Unfortunately, only starting material was obtained but the addition of another equivalent of DIBALH resulted in the formation of the alcohol. The reaction proceeded to completion and the pure product, **122**, was isolated by extractive workup, using Rochelle's workup solution to break the aluminium emulsion.

The desired aldehyde, **118** was then obtained by the Swern oxidation of **122**. The reaction proceeds to completion to give the desired aldehyde, which was isolated via a simple extractive workup. The crude aldehyde, **118** was used in enzymatic reactions directly after the workup to avoid issues of aldehyde oxidation during any purification steps.

4.2.2 Reactions with NCS variants

Based upon work presented in Chapter 3, three *Tf*NCS mutants, M97V, M97F and L76V were shown to be particularly promising in providing (1*S*'1'*R*)-THIQ products of the reactions between dopamine and 2-phenylpropanal, in improved yields and diastereomeric excesses compared with the wild-type enzyme. Therefore, it was hoped that similar results would be observed for reactions between **5** and **118** to give **119** (Scheme 4.6). The aldehyde, **118** was used crude and it was likely only one enantiomer of aldehyde would be accepted so two equivalents were added to the reaction mixture and analogous reactions conditions were used to those used with other α-methyl substituted aldehydes. Obtaining values for the diastereomeric ratios of the resultant products was straightforward as the two diastereomers, (1*S*,1'*R*) and (1*S*,1'*S*), are separable by RP-HPLC (method 1), as shown in Figure 4.2. A major diastereomeric product was observed, again indicating a kinetic resolution of the α-methyl substituted aldehyde substrate.



Scheme 4.6: *Tf*NCS catalysed reaction performed to give the THIQ scaffold towards generating 13-MeTHPB alkaloids. *Reaction conditions*: **5** (10 mM), **118** (10 mM), sodium ascorbate (5 mM), Δ33*Tf*NCS (0.2 mg mL^{-1}) in HEPES buffer (100 mM, pH 7.5) with acetonitrile (10% *v/v*) at 37 °C for 3 h, 250 μL scale reactions. Conversions were determined based on product formation compared to calibration curves of the products.

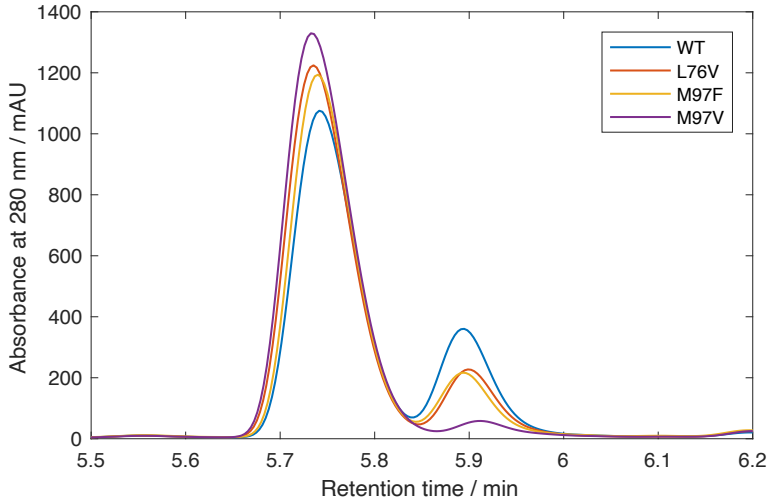


Figure 4.2: **Analytical HPLC analysis of product 119 formed by NCS-mediated Pictet-Spengler reactions.** Corresponding diastereomeric ratios and the reaction conditions used are given in Table 4.1.

No background reaction was observed, consistent with other NCS-mediated reactions with α-methyl substituted aldehydes, so it was believed that the product generated should have *S*-stereochemistry at the C-1 position. Based upon previous work (Chapter 3), it was rationalised that the major isomer formed was (1*S*,1'*R*) and the minor isomer was (1*S*,1'*S*).[148] The stereochemistry at C-1' of the product of NCS reaction between dopamine (**5**) and 2-phenylpropanal (**85**) has also been confirmed computationally to be (*R*), as increased near-attack conformations were observed between the two substrates.[148] To further confirm stereochemistry of this product, the structure would have to be determined by small molecule crystallography which would require significant amounts of material. Due to the highly acidic nature of the α-proton of **118**, the aldehyde is highly prone to racemisation, so performing enzymatic reactions with the two single enantiomer aldehydes is not feasible. The diastereomeric ratios of the products obtained, using four different *Tf*NCSs; wild-type, L76V, M97F and M97V, are given in Table 4.1.

Table 4.1: **Diastereomeric ratios of 119, generated by NCS-mediated reactions.** Reaction conditions were the same as those described for Scheme 4.6.

*Tf*NCS variant used	Diastereomeric ratio of **119** (1*S*,1'*R*):(1*S*,1'*S*)	Diastereomeric ratio of **96** (1*S*,1'*R*):(1*S*,1'*S*)
Wild type	77:23	94:6
L76V	87:13	97:3
M97F	87:13	94:6
M97V	96:4	97:3

NCS-catalysed reactions between dopamine (**5**) and 2-phenylpropanal (**85**) (to give **96**) using either mutant, M97V or L76V gave the highest diastereomeric ratios in **119** generated (97:3 1*S*,1'*R*:1*S*,1'*S*). M97F also had good selectivity. Reactions with the M97V variant were comparable with **118** as the aldehyde substrate, with a 96:4 diasteromeric ratio achieved. It is therefore interesting here, that with the addition of *meta*- and *para*-methoxy groups to the aldehyde ring, that a reduced diastereomeric ratio was observed when using the L76V variant. Consideration of the crystal structure with a non-productive mimic of the reaction between **5** and (*R*)-2-phenylpropanal (**85**) (PDB: 6RP3), the phenyl group of the aldehyde part of the mimic sits close to L76 (Figure 4.3). The mutation leucine to valine, creates more space in the active site enhancing the fit of the intermediate with (*R*)-**110**, to give a higher diastereomeric ratio in the product. However, the dimethoxylated aldehyde **118**, is bulkier and so activity is altered. For reactions with all NCSs tested, complete consumption of dopamine was observed, indicating complete conversions to give the products. This was confirmed by HPLC-based calibration curves of **119**, with all products formed in >99% yield. Based on these results, all enzymatic reactions to give **119** were performed using *Tf*NCS-M97V.

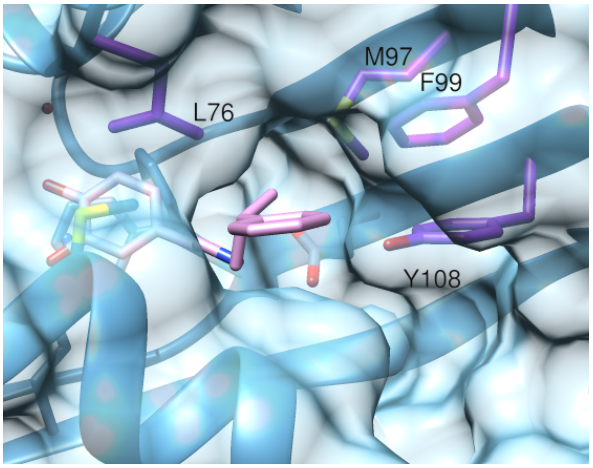


Figure 4.3: Structure of co-crystallised wild-type *Tf*NCS with (2*R*)-phenyl-propanal reaction mimic, (*R*)-110 bound in the active site. The phenyl ring of the aldehyde side chain is neighboured by L76, F99, Y108 and the α-helix residues A182 – G178.

The aldehyde, 2-phenyl-propanal (**85**) is known to be prone to racemisation at pH > 7, so it was expected that aldehyde, **118** could also racemise under the NCS reaction conditions at pH 7.5.[166] Therefore, the reaction was also explored using 1 equivalent of aldehyde, rather than 2 used and indeed analogous results were observed, with a diasteromeric ratio of 95:5 in the product observed with complete conversion (>99%) based upon a HPLC-based calibration curve of the product standard **119**.

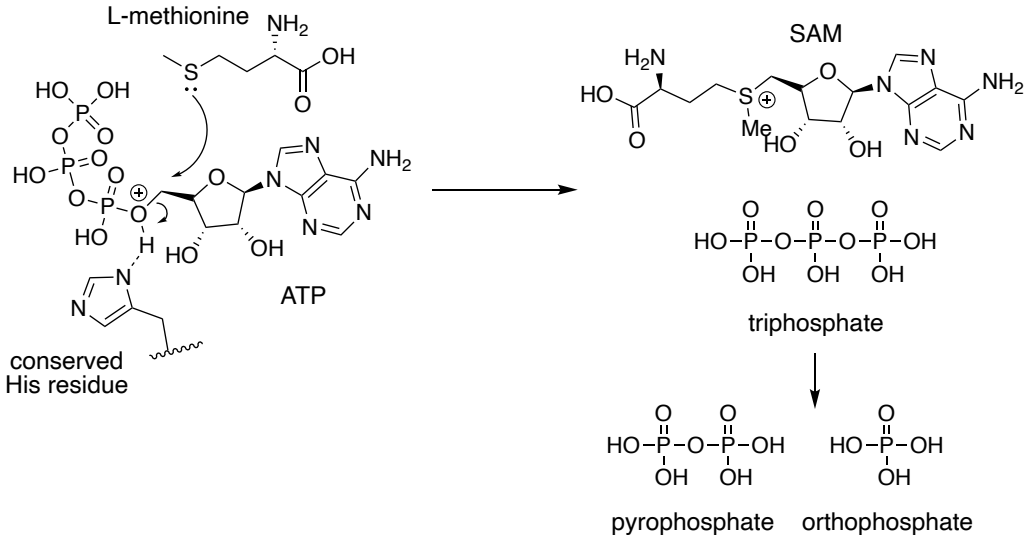
4.2.3 Increasing the molecular complexity of the THIQ scaffold using regioselective catechol *O*-methyltransferases

There are several research efforts into the discovery and isolation of methyltransferase enzymes, to avoid the challenges of synthetic, regioselective methylations.[201] These enzymes are used to selectively methylate a single heteroatom in a small molecule substrate. Regioselective methylation reactions are also particularly useful for drug discovery purposes due to 'the magic methyl effect' whereby the substitution of H to CH_3 can increase the potency of a drug molecule by up to 100-fold.[202] In 2011, it was noted that methyl groups are found in >67% of the year's top selling drugs.[202]

Beyond the advantage of controlled regioselectivity, biocatalytic reactions involving methyltransferases are more environmentally benign than traditional, chemical alternatives as they avoid the use of carcinogenic alkylating agents (for example, methyl iodide) as the methyl donor. However, the major barrier to the widespread applicability of these enzymes in preparative-scale syntheses is the narrow substrate scope of many of the enzymes discovered and the requirement for stoichiometric quantities of (*S*)-adenosyl-L-methionine (SAM). SAM is expensive (*ca.* £1000 for 1 g) and unstable (to hydrolysis), so reactions are not economically viable on a preparative scale. Recent efforts to overcome this have involved the use of an enzymatic co-factor generation system. Instead of using SAM directly, stoichiometric quantities of adenosine triphosphate (ATP) and L-methionine are added to the reaction with two other enzymes. SAM is then formed *in situ* using a methionine adenosyl transferase (MAT, E.C.2.5.1.6), which is used for methyl donation by the methyltransferases. This results in the formation of (*S*)-adenosyl-L-homocysteine (SAH), which can inhibit the methyltransferase reaction, so conversions are improved if SAH is degraded. A methylthioadenosine/SAH nucleosidase (MTAN, E.C.3.2.2.9) can therefore also be added to the reaction mixture to break down SAH, giving adenine and (*S*)-ribosyl-L-homocysteine.[203]

The MAT and MTAN enzymes investigated in this work to supply SAM and to degrade inhibitory SAH, respectively were based upon previous successful reports.[203] *Ec*MAT and *Ec*MTAN (from *Escherichia coli*) were available our group from a collaborator and were already under investigation by other group members (Fabiana Subrizi, Hailes group). *Ec*MAT is the most widely studied methionine adenosine transferase enzyme, and the mechanism of action has been elucidated with crystallographic studies.[204] The reaction proceeds via an S_N2 reaction between L-methionine and ATP. The triphosphate moiety of ATP is cleaved and SAM results. Triphosphate is then broken down into pyrophosphate and orthophosphate which are released with SAM by the

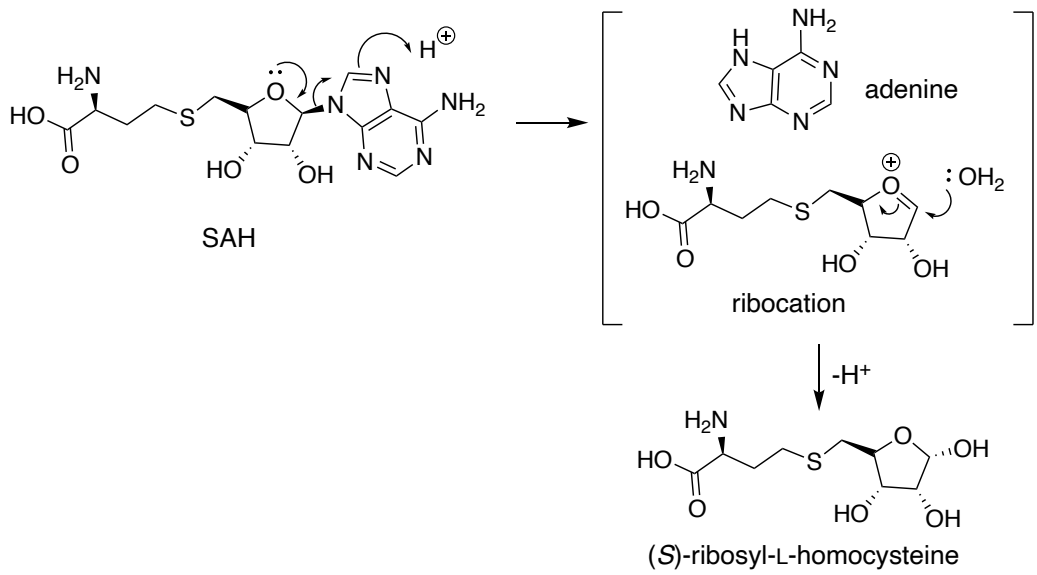
enzyme. *Ec*MAT is a dimeric enzyme, with two active sites at the oligomer interface. The mechanism has been reported, although elucidation faced significant challenges as multiple ligand conformations were observed in crystallographic studies. However, it has been hypothesised that a conserved histidine residue likely protonates the *O*-5' atom of ATP, providing a driving force for SAM formation. Substrate binding residues are conserved between other MAT enzymes isolated from both bacterial and mammalian sources (with one exception) indicating that the mechanism is conserved. A flexible loop (residues 98-114) close to the active site has also been noted and is thought to be implicated in the control of substrate binding.[205,206] Magnesium (Mg^{2+}) and potassium (K^{+}) ion co-factors are also required. The hypothesised mechanism of SAM formation by MAT enzymes is given in Scheme 4.7.



Scheme 4.7: Mechanism of SAM formation by MAT enzymes.

The *E. coli* derived SAH nucleosidase, *Ec*MTAN catalyses the cleavage of the glycosidic bond in SAH to generate adenine and (*S*)-ribosyl-L-homocysteine. In eukaryotes and plants, the catabolism of SAH generates adenosine and L-homocysteine instead.[207] *In vivo*, these enzymes act as feedback inhibitors to control biological methylation; when they are inhibited, levels of SAH increase and methyltransferases are inhibited. The enzyme is catalytically active when homodimeric

and proceeds via the formation of a ribocation to which there is nucleophilic addition of water to give (*S*)-ribosyl-L-homocysteine, as demonstrated in Scheme 4.8.[208]



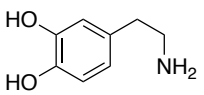
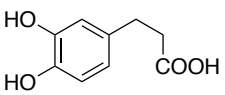
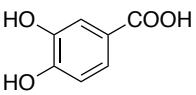
Scheme 4.8: Degradation of SAH by *Ec*MTAN.

Numerous regioselective catechol *O*-methyltransferases (COMTs) have been reported in the literature and shown to accept various catecholamines as substrates. Two known catechol *O*-methyltransferases; *Rn*COMT (isolated from *Rattus norvettigus*) and *Mx*SafC (isolated from *Myxococcus xanthus*) were explored with the 13-MeTHPB generated as the substrate.[209]

For methylations involving *Rn*COMT, methylation has been most often reported to occur on the *meta*-hydroxyl. The known substrate scope includes a variety of catecholamines (dopamine (**5**), dihydrocaffeic acid (**123**) and 3,4-dihydroxybenzoic acid (**124**))[209,210] and in our group, we have been exploring them with 1-benzyl THIQs as substrates.[197] The ionisation of the substrate is thought to modulate the observed methylation pattern.[210] The enzyme is often not completely regioselective, and typically some *para*-methylation is observed, for example with **5**. *Mx*SafC is reported to have complementary regioselectivity to *Rn*COMT, with methylation often favoured at the *para*-position with an exception being **124**, although few substrates have been

screened.[211] *Mx*SafC is involved in the biosynthesis of saframycin MX1, which possesses anti-tumour and antibiotic activites.[212] Catechol-containing molecules such as catechol, dopamine and caffeic acid are also accepted as substrates. The regioselectivity observed with *Mx*SafC is often higher than that observed with *Rn*COMT, however this is substrate dependent.[209] The regioselectivities of the products generated methylation reactions of three substrates for *Rn*COMT and *Mx*SafC are given in Table 4.2.

Table 4.2: **Observed regioselectivities of two catechol *O*-methyltransferases, *Rn*COMT and *Mx*SafC for reactions with three catecholamines**.

	 **5**	 **123**	 **124**
*Rn*COMT	83% *meta* 17% *para*	50% *meta* 50% *para*	59% *meta* 31% *para*
*Mx*SafC	<3% *meta* >97% *para*	5% *meta* 65% *para*	64% *meta* 7% *para*

The mechanism and regioselectivity of *Rn*COMT has been explored through active-site modification and crystallographic studies. A successful reaction requires the binding of Mg^{2+} to the active site before SAM, although there is no direct interaction between the two co-factors. The catechol binds to the magnesium ion, lowering the pK_a of the hydroxyl groups. The orientation of catechol binding determines the regioselectivity of the methylation as the hydroxyl group closest to the methyl group of SAM is deprotonated by K144, leading to methylation. Active site variants (notably W38D, W38R and Y200L) were able to improve regioselectivity towards either hydroxyl group.[213] A similar mechanism of action occurs with *Mx*SafC, with many key active site residues conserved.[211] The differences in observed regioselectivities between the catechol *O*-methyltransferases is subtle, with many factors at play beyond protein structural differences, including substrate characteristics and reaction conditions.[211]

To increase the complexity of the THIQs product (**119**) generated from an NCS reaction, towards analogues of corydaline (**90**), the two regioselective catechol *O*-methyltransferases, *Mx*SafC and *Rn*COMT were explored with **119** as the catechol-containing substrate. Expression of the two COMTs was performed using a published protocol, giving high expression of both enzymes[209] and biocatalytic reactions were attempted using clarified cell lysate as the reaction catalyst. Although reactions using *Rn*COMT were successful, no conversions were observed when using *Mx*SafC. It was suspected that this was due to issues of enzyme instability and indeed, protein precipitation was observed when the enzyme was purified by nickel-affinity chromatography. During dialysis to buffer exchange the protein sample, a significant amount of protein precipitation occurred, likely due to the incompatibility of the protein samples in high concentrations of imidazole even for a short period of time. Regardless, a high yield of pure protein was isolated (55 mg from a 500 mL culture). SDS-PAGE analysis of the expression of *Rn*COMT and the purification of *Mx*SafC, *Ec*MTAN and *Ec*MAT by Ni-affinity chromatography are given in Figure 4.4.

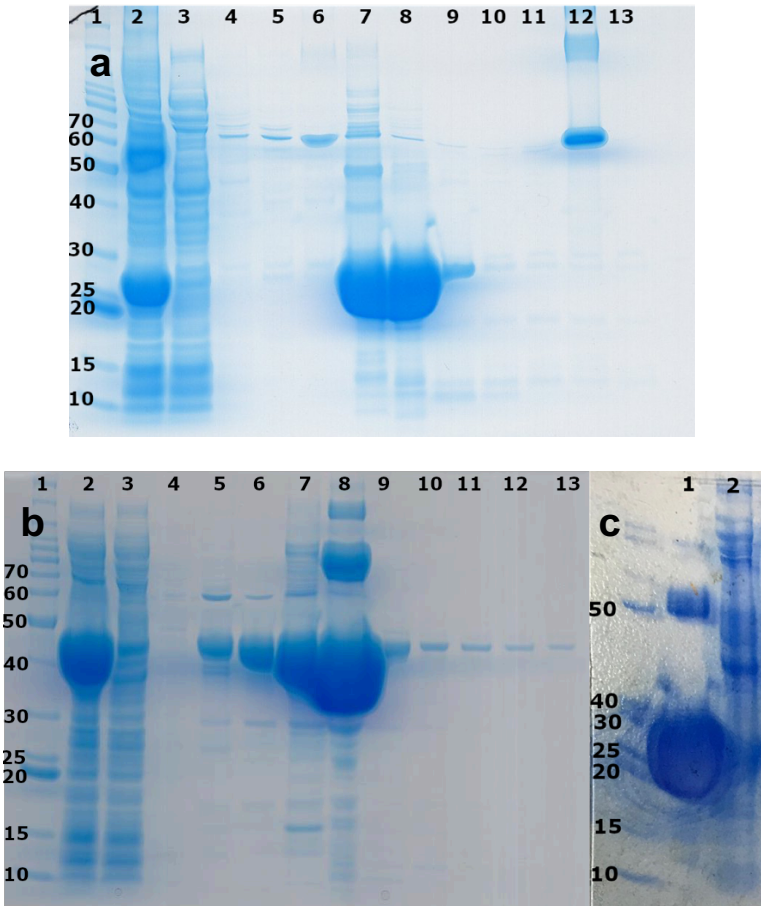
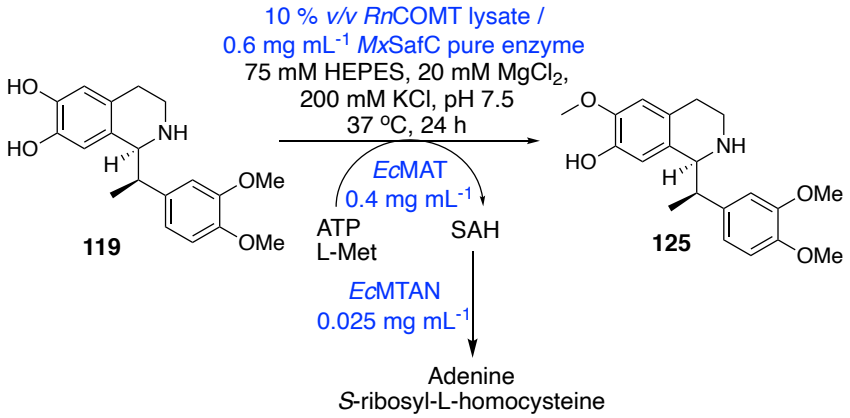


Figure 4.4: SDS-Page analysis of *Rn*COMT, *Mx*SafC, *Ec*MAT and *Ec*MTAN lysate or purifications. a. His-trap purification of *Mx*SafC. Lanes: 1, Benchmark™ Protein Ladder (masses given in kDa) 2, clarified cell lysate loaded onto column. 3, flow through with lysis buffer. 4, wash with 6 CV 20 mM imidazole buffer. 5, wash with 6 CV 40 mM imidazole buffer. 6-13, wash with 500 mM imidazole buffer and collected 3 mL fractions. **b**. His-trap purification of *Ec*MAT. Lanes: 1, Benchmark™ Protein Ladder (masses given in kDa) 2, clarified cell lysate loaded onto column. 3, flow through with lysis buffer. 4, wash with 6 CV 20 mM imidazole buffer. 5, wash with 6 CV 40 mM imidazole buffer. 6-13, wash with 500 mM imidazole buffer and collected 3 mL fractions. **c** Lanes: 0, Benchmark™ Protein Ladder (masses given in kDa). 1, purified *Ec*MTAN (some dimerization is observed), 2, expression of *Rn*COMT.

During initial screening experiments, both methyltransferases were found to be incompatible with the use of an organic co-solvent (at 10% *v/v*). Therefore, NCS reactions were performed using acetonitrile as the co-solvent rather than DMSO (as described for reactions with α-methyl substituted aldehydes in Chapter 3) as the boiling point is lower (82 ^{o}C vs. 189 ^{o}C). Once the NCS reaction was complete (monitored by RP-HPLC), the reaction mixture was quenched by addition of acetonitrile, to precipitate

the enzyme. The reaction mixture was then centrifuged, and the supernatant concentrated under reduced pressure or lyophilised to remove the water and acetonitrile. The resulting residue was then resuspended in buffer, containing all the metal co-factors required for a successful methyltransferase reaction, considering that the HEPES from the NCS reaction remained. This method also had the added benefit of not needing to isolate the THIQ intermediate.

Reactions were performed using the co-factor generation system detailed in Scheme 4.9 using purified enzyme as previously reported.[203] Conditions were developed in collaboration with Fabiana Subrizi (Hailes group) who was working towards the regioselective *O*-methylation of benzylic THIQs generated by NCS reactions,[197] and conditions used here involve increased concentrations of enzymes (methyltransferases, MAT and MTAN) and co-factors (ATP and L-methionine) with the aim of gaining complete conversions towards the products.



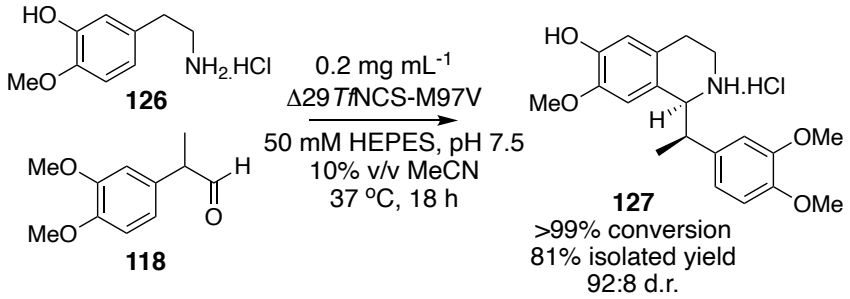
Scheme 4.9: The regioselective methylation reactions performed using catechol *O*-methyltransferases, *Rn*COMT or *Mx*SafC. 119 is methylated to give **125.** SAM recycling system enzymes are used to avoid the addition of stoichiometric quantities of the methyl donor, SAM. SAM is generated *in situ* by the enzyme, *Ec*MAT and the reaction byproduct, SAH is broken down by another enzyme, *Ec*MTAN.

Analytical HPLC analysis of the product (**125**) generated using either methyltransferase (*Rn*COMT or *Mx*SafC), compared traces with the starting material (**119**) showed that

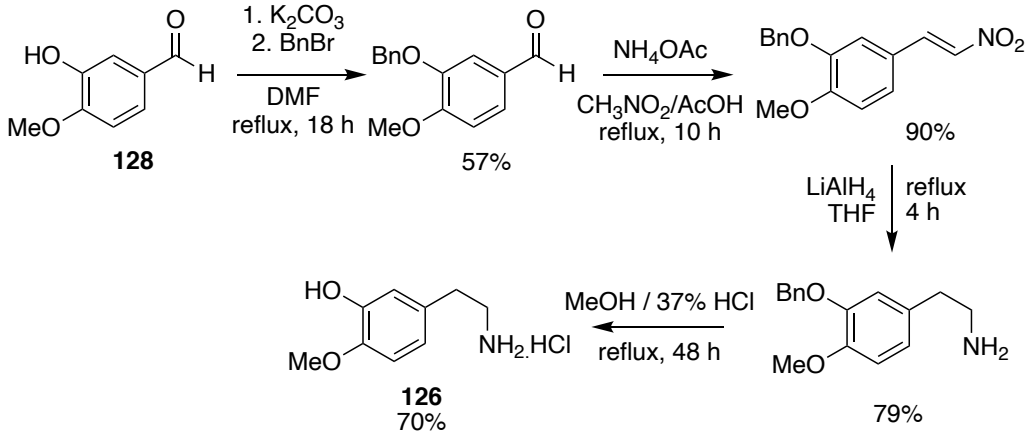
complete consumption of the starting material had occurred. Reactions were performed on a 20 mL, 5 mM scale and the resulting products isolated by preparative-HPLC (method 8) with the regioselectivity of the reaction determined by 2D-NMR spectroscopy (shown in the Appendix, Section 9.5). Using either methyltransferase, methylation was observed on the C6-position, with only a single isomer of **125** observed. High yields were also obtained (complete consumption of the starting material when using either *Rn*COMT or *Mx*SafC) and the product was obtained as the TFA salt in 12% isolated yield on a 5 mM, 20 mL scale using *Rn*COMT as the methyltransferase. This also follows literature precedent using *Rn*COMT, where methylation is only observed in the C6 position for reactions with **5**. This high regioselectivity is interesting as usually a mixture of the two regioisomers is generated (Table 4.2).[209] *Mx*SafC is reported to have a more promiscuous methylation pattern, but predominately reacts at the C7 position.[211] However, the differences in regioselectivity between reactions with **123** and **124** demonstrate the subtlety of the reaction outcome, so it was not completely unexpected that the opposing regioselectivity was observed with these THIQs as substrates.

4.2.4 Synthesis of C7-*O*-methoxy THIQs

Both catechol *O*-methyltransferases investigated, *Rn*COMT and *Mx*SafC were observed to be highly regioselective towards the C6-OH, so the C7-methoxylated isomer could not be generated in this way. Productive NCS reactions are observed with dopamine analogues lacking the C7-OH, so to generate C7-methoxylated analogues, NCS reactions were attempted between a dopamine analogue, **126** and the aldehyde, **118**, to give **126** (Scheme 4.10). **126** was synthesised by a MSci student, Joseph Broomfield (Hailes group) in four steps (28% overall yield) starting from isovanillin (**128**) as shown in Scheme 4.11.



Scheme 4.10: Reaction towards *para*-*O*-methoxy THIQs. *Reaction conditions*: **126** (10 mM), **118** (20 mM), sodium ascorbate (10 mM), purified Δ29*Tf*NCS-M97V (final concentration of 0.2 mg mL^{-1}) in HEPES buffer (50 mM, pH 7.5) with MeCN (10% *v/v*) at 37 °C for 18 h. 10 mL scale reactions.



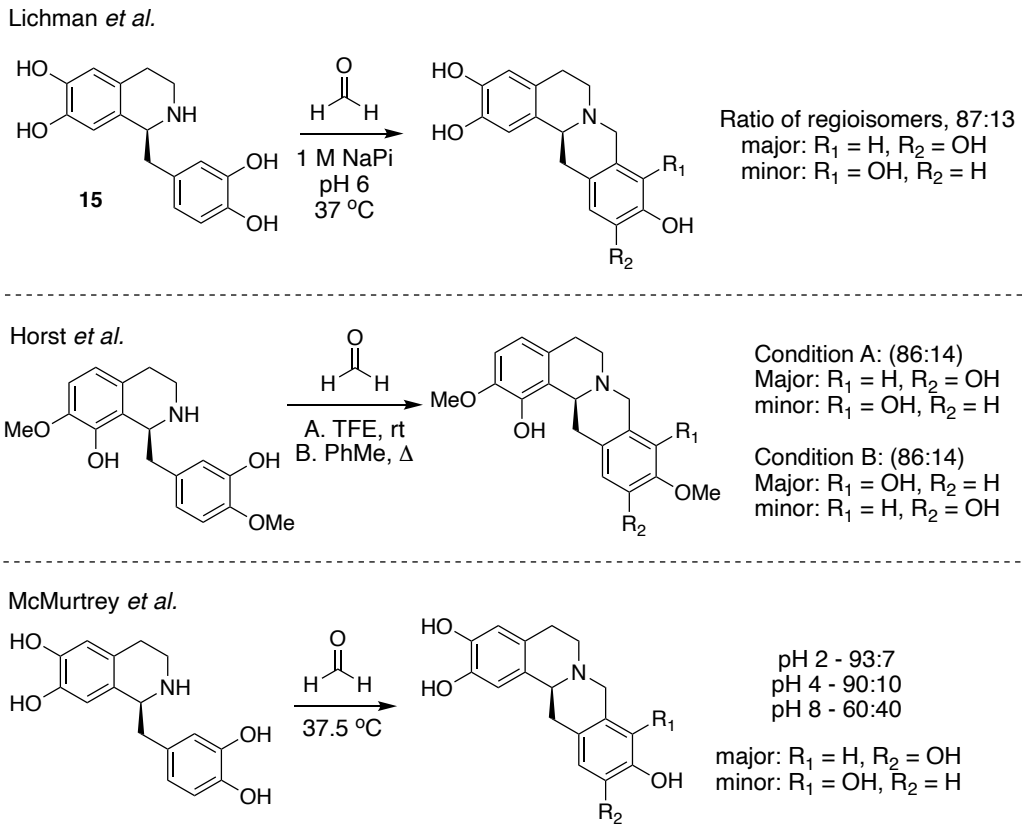
Scheme 4.11: Synthesis of 4-methoxytyramine (126). Reactions were performed by a Master's student, Joseph Broomfield (Hailes group).

The THIQ product (**127**) was isolated by preparative-HPLC (method 9) in >99% HPLC yield and 81% isolated yield with a 92:8 d.r. (1*S*,1'*R*:1*S*:1'*S*)-**127** as determined by ^{1}H-NMR spectroscopy (shown in the Appendix, Section 9.5). The slight decrease in selectivity compared with using dopamine as the phenethylamine substrate may be due to the lack of an additional hydrogen bonding interaction at the C7-OH with K122. The C6-OH of the catechol is essential for a productive NCS reaction and although the C7-OH is not, the lack of this interaction has been shown previously to lead to decreased conversions.[148]

4.2.5 Generation of the tetracyclic tetrahydroprotoberberine scaffold towards corydaline-like alkaloids

The formation of the tetracyclic protoberberine scaffold (Figure 4.1) from a benzylic THIQ requires a -CH_2- unit bridging the nitrogen, N-7, and the aromatic carbon, C-8a. In plants, the biosynthesis of the THPB, (*S*)-scoulerine (**11**) from a benzylic THIQ, involves a C-C bond forming reaction between an *N*-methyl group (on (*S*)-reticuline, **9**) and the benzylic moiety by the berberine bridge enzyme (BBE),[74,75] which also provides regioselectivity in the ring closure step, leading to methoxy and hydroxyl substituents on the D-ring at positions 9 and 10 (Scheme 4.1). The substrate scope of the BBE has not been widely explored, so chemical alternatives to generate the tetracyclic scaffold were instead considered.

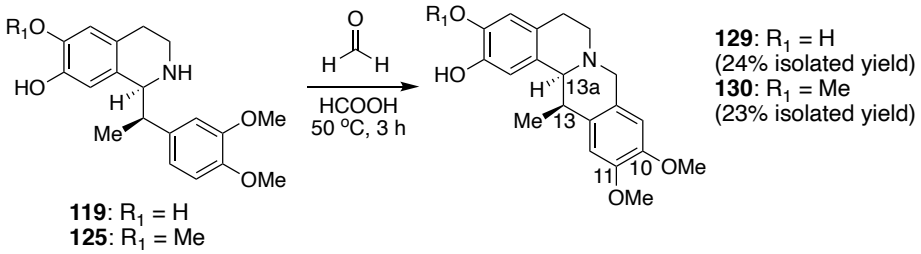
Intramolecular, chemical Pictet-Spengler reactions on THIQs have proved successful in generating a range of berberine-like alkaloids. Lichman *et al*. used a phosphate-catalysed Pictet-Spengler reaction between norlaudanosaline (**15**) and formaldehyde to generate two THPB alkaloids with 10,11-dihydroxy substitutions on the D ring i.e. with complementary regioselectivity to the BBE route (Scheme 4.12).[117] The effect of solvent on the regioselectivity of the Pictet-Spengler reaction has been explored by Horst *et al*. where a range of *ortho*-hydroxylated THPBs were generated (Scheme 4.12).[198] The regioselectivity of the Pictet-Spengler reaction (between the THIQ and formaldehyde) upon C-ring formation was shown to be dependent on whether the reaction solvent used was protic or aprotic. Using aprotic solvents resulted in 9,10 substitution on the D-ring, whereas using protic solvents results in complementary regioselectivity. Studies of the Pictet-Spengler between norlaudanosaline (**15**) and formaldehyde in aqueous buffer performed by McMurtrey *et al*. have also indicated that pH can also influence regioselectivities.[214] The three routes previously explored with the products generated are summarised in Scheme 4.12.



Scheme 4.12: Chemical Pictet-Spengler reactions between THIQs and formaldehyde towards the THPB scaffold.

It was therefore hoped that these routes could be extended towards generating a range of 13-MeTHPBs with different regioselectivities. Initially, Pictet-Spengler reactions between **119** and formaldehyde were attempted using crude, lyophilised NCS reaction products with analogous conditions to those used by Lichman *et al*.: formaldehyde in 1 M sodium phosphate buffer (pH 6) at 37 °C. Unfortunately, no desired product was observed by LCMS after 30 min (as reported by Lichman *et al*.) nor after longer reaction times (24 h) or with heating of the reaction mixture to 60 °C. Compared with the THIQ substrate used by Lichman *et al*., **119** has methoxy substitution on the D-ring rather than hydroxyl groups. Methoxy groups are less activating than hydroxyl groups and this therefore suggested that more forcing conditions would be required. Therefore initially, acid-catalysed Pictet-Spengler reactions were attempted in protic solvents. No reaction was observed using either methanol or 2,2,2-trifluroethanol as the reaction solvent. Therefore, analogous conditions were used to those performed by Qian *et al*. in the

synthetic development of a range of THPB alkaloids, using formaldehyde and formic acid as the reaction solvent and heating to 50 °C. Compound **125** was initially used as the starting material, as it was less oxidatively sensitive as the catechol was protected by the methyl group. Complete consumption of the THIQ starting material, **125** was observed by RP-HPLC after 2 h and subsequent NMR and MS analyses determined that the resultant product was **130** (Scheme 4.13) with a 10,11-dimethoxy substitution pattern on the D-ring. THPB **130** was isolated in 23% yield after purification by preparative-HPLC (method 9). The reaction also went to completion when using the analogous, non-methoxylated THIQ **119** as the starting material and the product (**129**) was isolated in 24% yield. The same regioselectivity in the D-ring was observed. As for **129**, Nuclear Overhauser Effect Spectroscopy (NOEsy) NMR analysis of the two products isolated confirmed the stereochemistry at C-13 and C-13a (see Appendix, Section 9.5). A strong NOE correlation was observed between 13-H and 13a-H, indicating *syn* geometry. A very weak NOE signal was also observed between 13a-H and the methyl group protons at C-13, implying *anti* stereochemistry. (*S*)-Stereochemistry at C-13a is most likely to be generated from the NCS reaction, the stereochemistry at C-13 must therefore be (*R*), further confirming previous experimental results on the absolute stereochemistry at these positions.[148]



Scheme 4.13: General route of the chemical Pictet-Spengler reactions performed between formaldehyde and THIQs 119 and 125.

Formation of the other regioisomeric product i.e., 9,10-dimethoxy substitution on the D-ring, has been reported by exploiting solvent effects.[198] In theory, if an apolar, aprotic reaction solvent is used, the iminium ion intermediate of the reaction can be stabilised

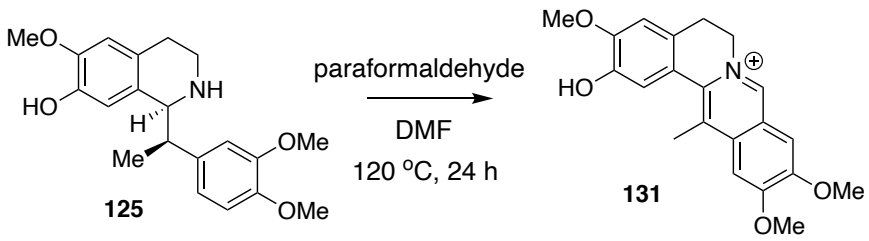
by an electrostatic interaction between the iminium ion and a hydroxyl group on the benzylic moiety. This holds the iminium ion intermediate in an orientation consistent with product formation with a 9,10-dimethoxy substitution pattern. Whereas, when a protic solvent is used the intermediate is stabilised by the solvent, leading to the less hindered product being formed, with 10,11-dimethoxy substitution.[198]

To attempt to generate the complementary regioisomer to **129** and **130** in the D-ring, i.e. the 9,10-dimethoxy product, Pictet-Spengler reactions were attempted using apolar, aprotic reaction solvents following methods developed by Horst *et al*.[198] (detailed in Scheme 4.12). Paraformaldehyde was used instead of formaldehyde (which is commercially available at 37% in water) to ensure that no proton sources were present in the reaction mixture and reactions were also performed under anhydrous conditions. The starting material was furthermore prepared as the free amine to aid solubility. As reported by Horst *et al*., the best regioselectivity towards products with 9,10-dimethoxy substitution, was observed when using toluene as the reaction solvent (81:19 ratio of 9,10:10,11 dimethoxy substitution), with heating at 105 ^{o}C for 4 days (Scheme 4.12). Identical reaction conditions were attempted here, however there was only trace evidence of product formation by LCMS analysis. Low conversions were likely to due to poor solubility of the starting material in toluene. Reactions were also attempted using 1,2-dichloroethane as the reaction solvent however no product was observed. It is likely that the presence of two electron donating methoxy groups on the benzylic moiety means the pi-system is insufficiently activated. Therefore, more forcing conditions may be required, explaining why formic acid catalysed reactions were successful, and reactions performed with just heating in apolar, aprotic solvents were not.

4.2.6 Syntheses of 13-methylprotoberberine alkaloids

Reactions with other aprotic, high-boiling point solvents were considered to see if improved reactivities could be observed. The Pictet-Spengler reaction between **125** and

formaldehyde was attempted using DMF (*N,N*-dimethylformamide) as the reaction solvent with increased heating (120 °C) (Scheme 4.14). After 24 h, LCMS and RP-HPLC analyses showed complete consumption of the starting material (**125**) and identified a product with a *m/z* of 352 Da i.e. 3 Da lower than expected. This therefore led to the hypothesis that the oxidation of one of the rings in the tetracyclic scaffold had occurred. The resultant product was purified by preparative-HPLC analysis (with an isolated yield of 5% and an HPLC yield of >99%) and 2D NMR analysis was used to determine the structure as **131** (Scheme 4.14). NMR spectra are given in the Appendix, Section 9.5.

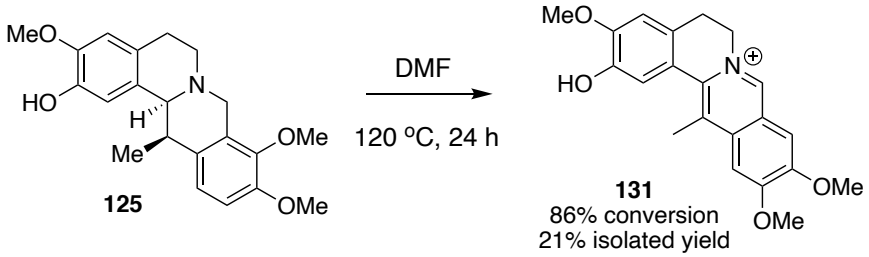


Scheme 4.14: Chemical Pictet-Spengler reaction between 125 and formaldehyde using DMF as the reaction solvent.

The presence of four singlets in the aromatic region suggested that the 10,11 regioisomer of product was formed and the absence of the proton at position 13, indicated that the C-ring had been oxidised, giving **131**. Although removing the chirality at C-13 and C-13a, this reaction may provide a useful route to generating protoberberine analogues of THPB alkaloids as the oxidation of the C-ring must occur after the Pictet-Spengler reaction.

Therefore, to determine whether protoberberine alkaloids could be generated by simply heating the correspond THPB in DMF, **125** was isolated and heated in DMF at 120 °C for 24 h (Scheme 4.15). The product was formed in 86% conversion by starting material depletion calibration curve and identity confirmed by ^{1}H-NMR and LCMS. An isolated yield of 21% was achieved after purification by preparative HPLC. This is a useful route towards protoberberine alkaloids directly from the corresponding THPB, avoiding the

use of a selective oxidase enzyme such as (*S*)-tetrahydroprotoberberine oxidase (STOX)[215] or harsh reaction conditions (I_2 / EtOH, reflux).[216] Since the stereochemistry at C-13 and C-13a is lost during the oxidation, the THIQ starting material could also instead be generated via a phosphate-mediated Pictet-Spengler reaction rather than an NCS reaction.[15]



Scheme 4.15: Single step reaction towards 13-methylprotoberberine, 131 from the analogous 13-MeTHPB 125.

Routes towards these compounds are useful as a range of 13-MeTHPBs isolated from the methanolic extracts of *Corydalis saxicola* (**132-134**, Figure 4.5) have been shown to possess anti-hepatitis B activities *in vitro*.[184] The methyl group at C-13 had a notable effected on activity, as other THPBs screened lacking this methyl group were inactive. These compounds are also reported to have sedative, anti-viral and anti-bacterial properties and routes towards the isolation of these compounds from *C. saxicola* are patented.[217] Protoberberine alkaloids can also be further functionalised at C-8 as the iminium ion is susceptible to nucleophilic attack.[218]

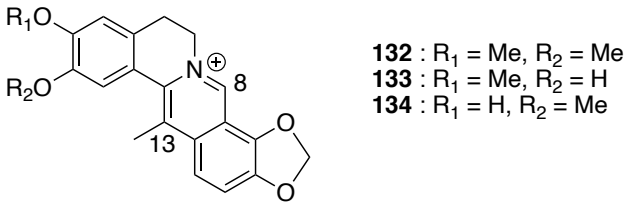


Figure 4.5: 13-Methylprotoberberine alkaloids possessing anti-hepatitis B activities.

4.2.7 Chemical Pictet-Spengler reactions towards C8-substituted THPBs

Two alkaloids isolated from *Corydalis ochotensis,* corytenchirine (**135**) and lienkonine (**136**) have been reported to possess methyl groups at C-8 instead of C-13 (Figure

4.6).[219,220] The stereochemistry of the C-8 centre in corytenchirine (**135**) was assigned based upon Bohlmann bands (2800-2650 cm^{-1}) whereby the C-H bond that is anti-periplanar to the nitrogen lone pair is considered to be elongated.[221] There are few examples of synthetic routes towards 8-MeTHPBs, with one stereoselective route reported involving the reduction of a 8-Me-protoberberine scaffold.[222] THPB alkaloids with methyl groups at both C-8 and C-13 have also been synthesised by Wiegrebe *et al*.[223] although no stereocontrol was achieved. One berberine-like alkaloid with a methyl group at C-8, coralyne (**137**) (Figure 4.6) is known to act as an anti-leukemic agent.[224] Therefore, routes towards similar alkaloids are useful for drug discovery purposes.

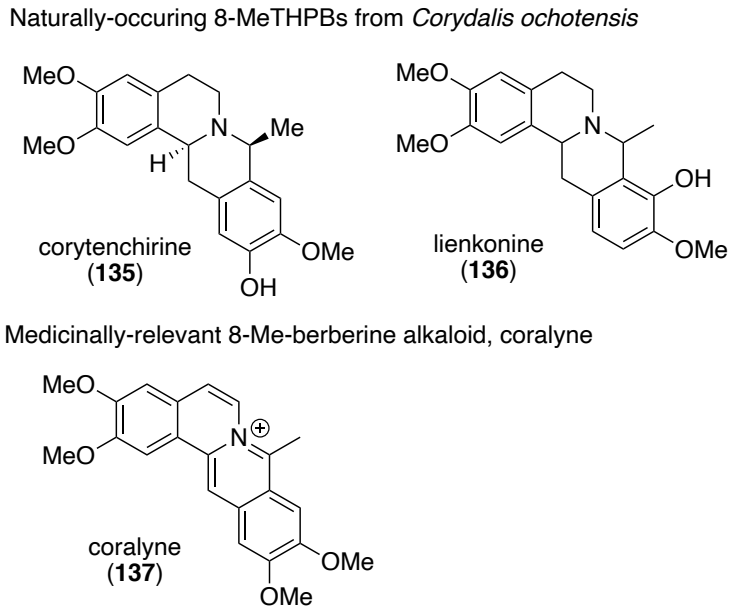
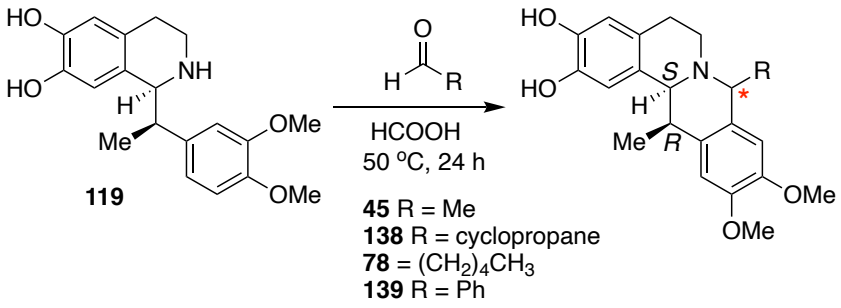


Figure 4.6: THPB and PB alkaloids with possessing a methyl group at C-8.

Therefore, to develop routes towards C-13, C-8 substituted THPBs, chemical Pictet-Spengler reactions were attempted between the NCS-generated THIQ (**119**) and a range of aldehydes (Scheme 4.16). Performing these reactions with aldehydes other than formaldehyde results in the formation of a new chiral centre at C-8 so it was also thought to be interesting to see whether the methyl group at C-13 would have an influence on this newly formed chiral centre. Acetaldehyde (**45**) was chosen due to the small size and cyclopropanecarboxaldehyde (**138**), hexanal (**78**) and benzaldehyde

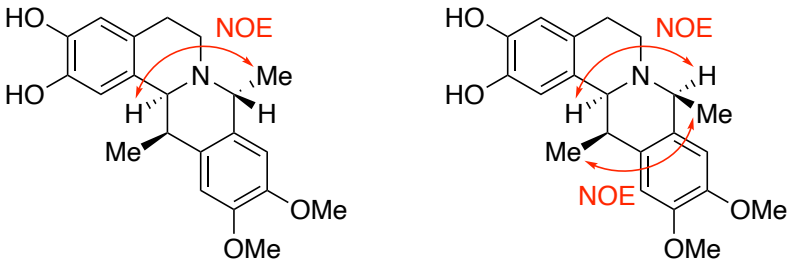
(**139**) were chosen to probe whether even bulkier aldehydes would be converted. Reactions were performed in formic acid as described for the reaction with formaldehyde and for an extended reaction time (24 h vs. 3 h with formaldehyde) as it was assumed that these reactions would be more challenging. The aldehyde was used in significant excess to help drive the reaction and all aldehydes could be easily removed under reduced pressure after the reaction was complete.



Scheme 4.16: Formic acid catalysed Pictet-Spengler reactions performed between 119 and four different aldehydes. A new chiral centre at C8 is formed during the reaction, giving three stereocentres in the THPB products.

Reaction mixtures were analysed by LCMS to detect for product formation. Product was only detected for the reaction of **119** with acetaldehyde (**45**), which was purified by preparative HPLC (56% conversion, determined by HPLC based upon starting material depletion). giving a 21% isolated yield. The product generated has three chiral centres; the stereochemistry at C-13 and C-13a is predetermined by the NCS reaction and the formic acid catalysed Pictet-Spengler reaction should not racemise these two chiral centres. NMR spectroscopic analysis of the product isolated (Section 9.5) indicated that the reaction was regioselective, with 10,11-dimethoxy substitution on the D-ring. Two epimers of product were present in a ratio of 3:1. The stereochemistry at C-8 could be determined by NOEsy because the stereochemistry at C-13a is known. For the major isomer, a NOEsy correlation was observed between the methyl group at C-8 and the proton at C-13a, indicating that the methyl group was below the ring, as drawn (Figure 4.7). The major isomer of product formed was therefore assigned as (8*R*,13*R*,13a*S*)-

140. For the minor isomer, (8*S*,13*R*,13a*S*)-**140**, the NOEsy spectrum should show a correlation between 8-H and 13a-H, however this could not be observed as the 13a-H proton was obscured by the residual solvent peak for water. A NOE signal could however be observed between the two methyl groups at C-13 and C-8 (shown in the Appendix, Section 9.5).



(8*R*,13*R*,13a*S*)-**140**
major isomer

(8*S*,13*R*,13a*S*)-**140**
minor isomer

Figure 4.7: Expected NOEsy correlations for the two isomers of 140 generated.

To rationalise why this isomer was predominately formed, energy minimised, 3D models of both epimers were generated (Figure 4.8). For both epimers, the C-ring is in a half-chair conformation. For the minor isomer generated, the methyl groups at C-8 and C-13 are on the same side of the ring, in a pseudoaxial orientation. This unfavourable steric interaction likely explains why formation of this isomer is disfavoured.

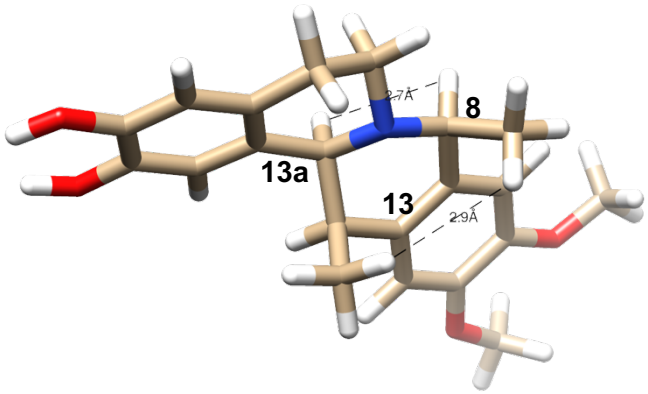
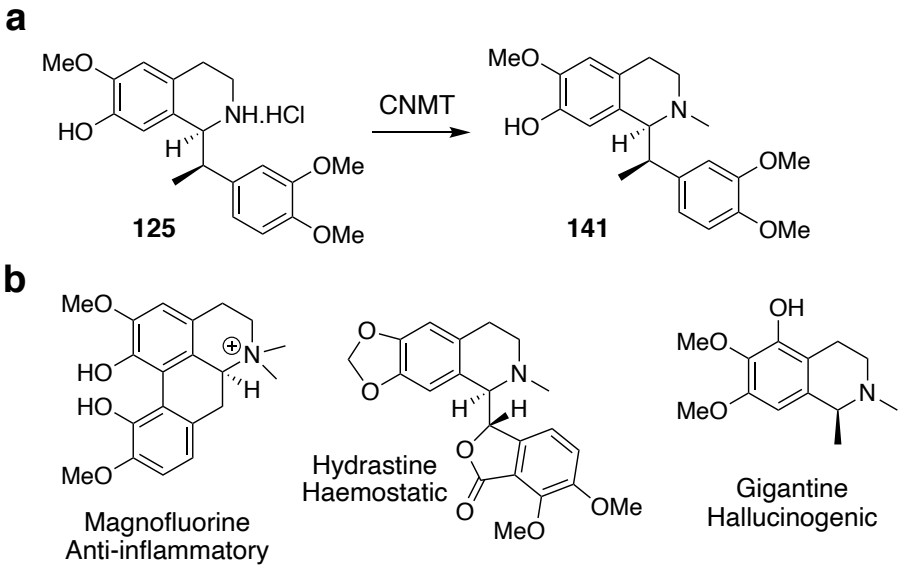


Figure 4.8: 3D modelling of the minor epimer formed, (8*S*,13*R*,13a*S*)-140. The two methyl groups at C-8 and C-13 are *syn* with a distance between the methyl group protons of 2.9 Å. The distance between 8-H and 13a-H is 2.7 Å.

Clearly extending the aldehyde substrate scope of the chemical Pictet-Spengler reaction between 1'-MeTHIQs such as **119** towards the analogous 13-MeTHPBs is challenging. One limiting factor could be the high volatility of the aldehyde and the requirement for forcing reaction conditions. Acetaldehyde has a low boiling point (20 ^{o}C) and heating to 50 ^{o}C was required for a successful reaction. To improve conversions towards completion, further equivalents of acetaldehyde could be added to the reaction mixture over time. It may therefore also be possible to incorporate longer chain, aliphatic aldehydes by using longer reaction times or using more forcing conditions, perhaps the use of a microwave.

4.2.8 Routes towards *N*-methylated 13-MeTHPBs

To further diversify the products generated, it was hoped that a regioselective *N*-methyltransferase could be used to selectivity mono-*N*-methylate the THIQ, **125** (Scheme 4.17), giving **141.** The *N*-methylated THIQ scaffold is a valuable pharmacophore, found in a range of pharmaceuticals and performing regioselective mono-*N*-methylations is challenging chemically due to increased nucleophilicity of the nitrogen lone pair after each methylation.



Scheme 4.17: The generation and relevance of *N*-methylated THIQs. a. *N*-methylated analogues of NCS products generated could be accessed using a selective *N*-methyltransferase. **b.** A range of biologically active *N*-methylated THIQs.

One appropriate enzyme, Coclaurine *N*-methyltransferase isolated from *Coptis japonica* (*Cj*CNMT) was available in the laboratory, provided as a pure, lyophilised sample by Michael Richter. The *in vivo* role of CNMT is to selectivity mono *N*-methylate (*S*)-coclaurine (**8**). Previous studies with *Cj*CNMT have shown that the enzyme has a broad substrate scope, although the THIQ scaffold appears to be essential for a productive reaction. The enzyme is also not stereospecific, accepting both enantiomers of the natural substrate, **8**. A range of THIQs accepted as substrates are given in Figure 4.9,[225] although the enzyme has low catalytic activity and complete conversion has not yet been observed with any substrate. The enzyme can accommodate THIQ with a range of substituents at C-1, including two racemic 1'-substituted THIQs, **142** (5% conversion) and **143** (87% conversion).[226] Therefore, it was hoped that this could be extended towards using the (1*S*,1'*R*)-THIQ,

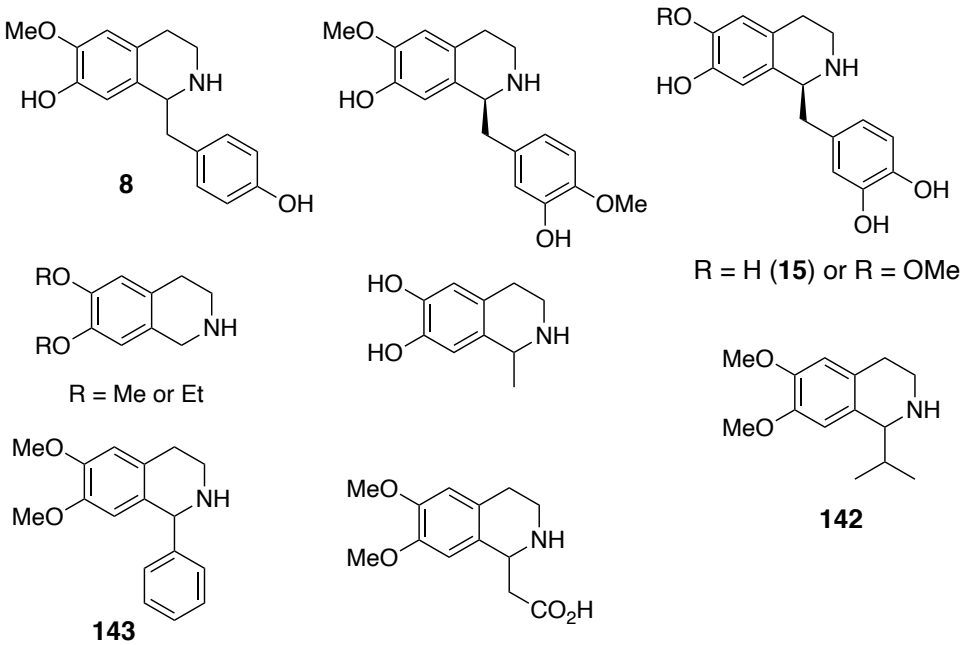


Figure 4.9: **A range of THIQs accepted as substrates by CNMT**, investigated by Choi *et al*.[225] and Bennett *et al*.[226] A range of THIQs beyond the natural substrate, (*S*)-coclaurine, are accepted as substrates.

The enzyme is SAM-dependent as the methyl donor for biotransformations, so enzyme reactions were attempted using the same co-factor generation system used with the

catechol *O*-methyltransferases (*Rn*COMT and *Mx*SafC). Enzyme reactions were again attempted with an excess of SAM formed *in vitro*, using 2 equivalents of ATP and L-methionine with **125** used as a crude product from the previous *O*-methylation reaction using *Rn*COMT. The lyophilised sample of CNMT was resuspended in HEPES buffer (50 mM, pH 7.5) and added to the reaction sample to give a final enzyme concentration of 0.5 mg mL^{-1} which was added once the *Rn*COMT reaction was deemed complete by analytical HPLC (method 2). Additional co-factors (1.5 e.q. of ATP and L-methionine) and SAM supply enzymes were added. LCMS and analytical HPLC (method 2) of the reaction mixture was used to determine if product formation was occurring.

Unfortunately, no product was observed, even with varying concentrations of CNMT (0.5, 1.0 and 2.0 mg mL^{-1} final concentration) and co-factors (5 mM, 10 mM, 20 mM), and increasing concentrations of the co-factor recycling system enzymes. Reactions were also attempted in a one-step reaction, adding both methyltransferases (CNMT and *Rn*COMT) at the same time, however no desired product was detected. The lack of acceptance of this substrate, **125** could be due to issues with not using a purified substrate. To generate enough SAM, 3 equivalents of ATP and L-methionine at least are required and perhaps the enzyme is incompatible with high phosphate concentrations. The remaining aldehyde, **2** from the NCS reaction may also interfere with the enzyme. To determine if this is the case, reactions could be attempted using purified, rather than crude starting material. It is also possible that this substrate is simply not accepted. Poor conversions (5%) were observed with another other 1'-MeTHIQ, **142**.[226] **125** is significantly larger and simply may not fit into the enzyme active site.

Enzymatic routes towards selective mono-*N*-methylation were unsuccessful so instead chemical routes towards di-*N*-methylated THIQs were explored. This scaffold is present in one naturally-occurring BIA, (*S*)-magnofluorine, an aporphine alkaloid which has

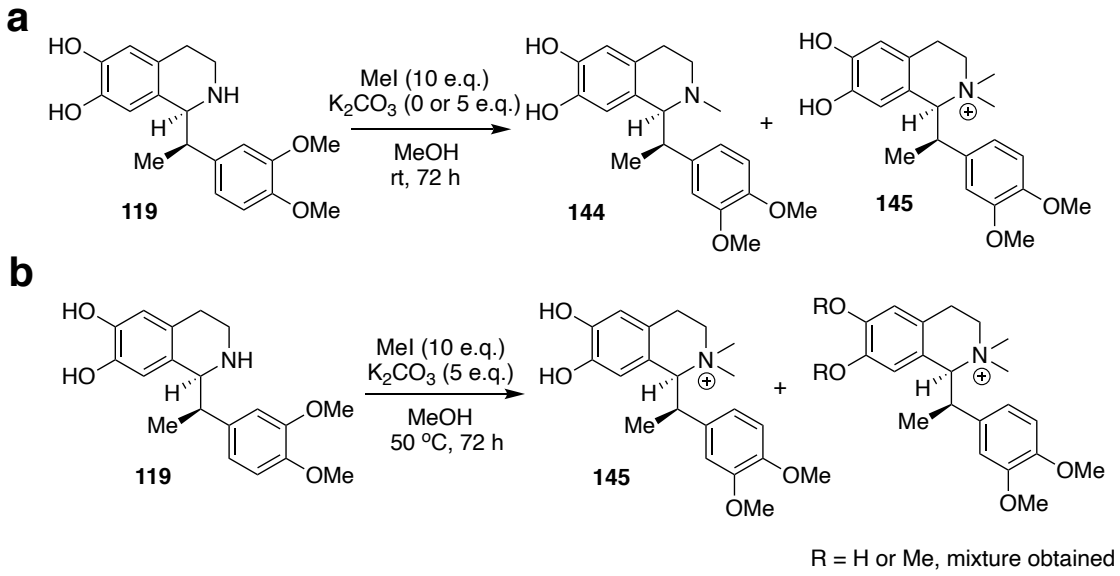
been shown to have anti-inflammatory, sedative and anti-fungal activities.[227,228] Quaternary nitrogen atoms are present in other pharmaceutically-relevant BIAs and 13-MeTHPBs (Scheme 4.17), so routes towards analogues are useful.

It was hoped that chemical conditions could be found which allow selective *N*-methylation whilst avoiding reactivity at the free hydroxyl groups to avoid a protecting-group based strategy. Trimethylammonium dopamine and trimethylammonium serotonin have been synthesised by Collins *et al*.[229] without protection of the free hydroxyl groups, using iodomethane as the methylating agent in excess (10 e.q.) with a single equivalent of potassium carbonate to deprotonate the nitrogen. The reaction was performed using crude substrate **4** with and without base as potassium carbonate is capable of deprotonating catechol hydroxyl groups (as seen during the benzyl deprotection of *N*-Boc-dopamine, Section 2.2.4). Reactions were performed under anhydrous conditions as iodomethane is easily hydrolysed to methanol.

For the reaction without base present, a mixture of starting material, mono- and di-*N*-methylated product (**144** and **145**) was observed by NMR spectroscopy. Increasing the number of equivalents of iodomethane used did not improve the selectivity. For reactions with base, a mixture of mono and di-*N*-methylated product was observed (Scheme 4.18a). Again, increasing the number of equivalents of iodomethane did not improve the outcome of the reaction. It was considered that more forcing reaction conditions would be required as perhaps, despite improved nucleophilicity, the mono-methylated nitrogen was sterically hindered. Therefore, 5 equivalents of potassium carbonate were used with heating to 50 ^{o}C. Although the quaternary nitrogen was fully formed, *O*-methylation at the *para*-position was also observed (Scheme 4.18b).

One-step chemical routes towards *N*-methylated THIQs are not straightforward, however with the constant discovery of novel enzymes, in future analogous routes could

be performed. **141** could also be generated via a synthetic route involving a protecting group strategy (i.e. via Boc-protection of the nitrogen, benzyl protection of the phenolic hydroxyl, *N*-methylation with iodomethane, followed by orthogonal protecting group removal under acidic conditions).



Scheme 4.18: Reactions performed towards selective *N*-methylations of 1'-MeTHIQs

4.3 Conclusions

A range of novel 13-MeTHPB alkaloids were generated, with opposing stereochemistry at C-13 and C-13a to the naturally occurring alkaloids isolated from *Corydalis plants.* This was carried out by extending work on the acceptance of α-methyl substituted aldehydes by norcoclaurine synthase in which (1*S*,1'*R*)-THIQs can be generated in high yields and selectivity in a single reaction step. Two regioselective catechol *O*-methyltransferases, *Rn*COMT and *Mx*SafC were shown to be completely regioselective towards the *meta*-hydroxyl of the THIQ generated, expanding the known substrate scopes of these enzymes, and providing a benign, facile, regioselective route towards the product. The methyl donor co-factor, SAM was generated using a MAT enzyme, to reduce the expense of the reaction, allowing for preparative scale reactions. The side-product, SAH formed after the methyltransferase reaction was degraded by a MTAN enzyme to avoid methyltransferase inhibition. This was based upon previous SAM

generation methods by Siegrist *et al*.[209] Chemical Pictet-Spengler reactions were then explored to generate the tetracyclic 13-MeTHPB scaffold. Due to reduced nucleophilicity of the aromatic D ring due to di-methoxy-substitution rather than phenolic groups, harsh reaction conditions involving formic acid were required to form the ring. Reactions involving formaldehyde or acetaldehyde as the aldehyde reagent were successful, both generating products with 10,11-dimethoxy substitutions on the D-ring. A new chiral centre was generated in the reaction with acetaldehyde, two epimers of product formed in a ratio of 3:1. Chemical Pictet-Spengler reactions between a 1'-Me-THIQ were also explored using DMF as the reaction solvent and interestingly, after the Pictet-Spengler reaction occurred, oxidation of the D-ring was observed, leading to the formation of the biologically relevant protoberberine scaffold. This reaction therefore provides a facile route towards berberine-like alkaloids from the corresponding THPB.

Chapter 5: Single Step Syntheses of (1*S*)-Aryl-Tetrahydroisoquinolines by Norcoclaurine Synthases

5.1 Introduction

5.1.1 Pharmaceutical relevance of 1-aryl-tetrahydroisoquinolines

The tetrahydroisoquinoline (THIQ) moiety is present in many biologically active molecules so is a desirable synthetic target to generate for pharmaceutical research. Generating single enantiomer and single regioisomer products is synthetically challenging. The 1-aryl-THIQ moiety is found in many synthetic biologically active small molecules (as given in Figure 5.1) which have been shown to have a variety of medicinal benefits, including anti-tumour[230], anti-HIV[231] and contraceptive[232] activities. Solifenacin succinate (**81**) is a widely-prescribed muscarinic-receptor antagonist, used for the treatment of an overactive bladder with around 3 million US prescriptions per year.[55]

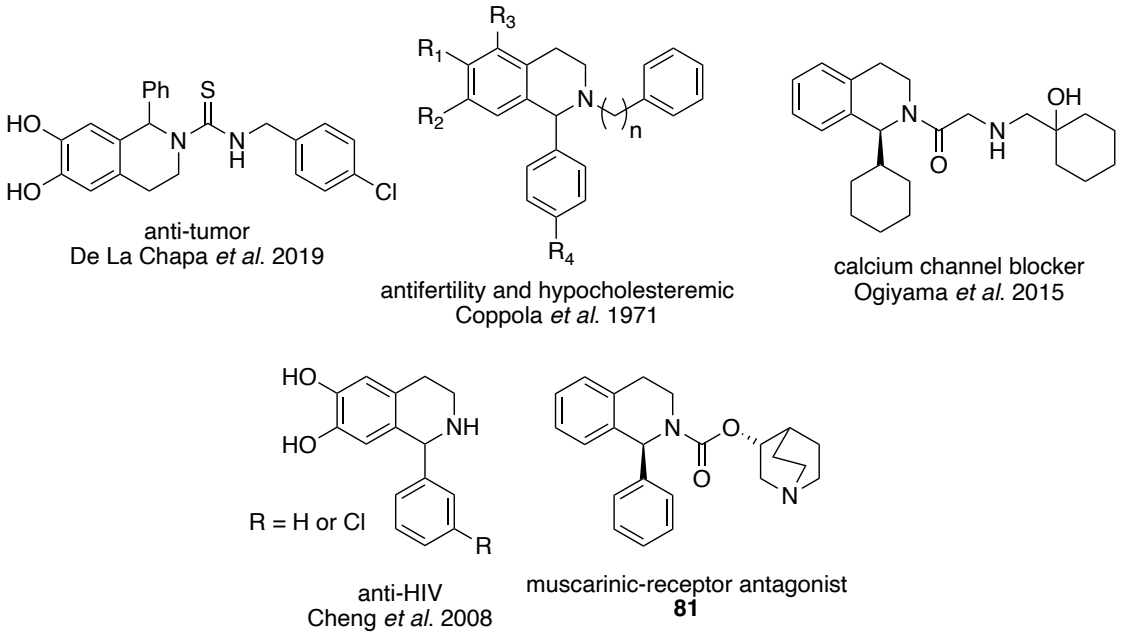


Figure 5.1: **A range of pharmaceutically relevant 1-aryl-tetrahydroisoquinolines.** Figure adapted from Roddan *et al.*[149]

The general route to racemic 1-aryl THIQs requires a three-step synthesis involving amide formation, a Bischler-Napieralski cyclisation to give the imine under forcing

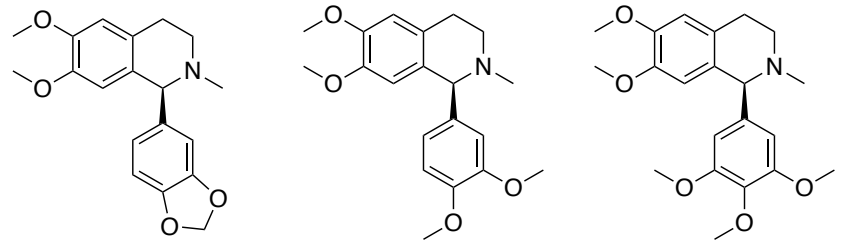
conditions, followed by reduction of the imine to give the final THIQ scaffold. Single-enantiomer syntheses of these molecules involve a stereoselective reduction of the imine using an inorganic catalyst (often iridium-based).[233–235] These catalysts however can have narrow substrate scopes and so are not a general, versatile route to generate a range of these compounds. In another synthetic route developed by Chen *et al.,*[235] the asymmetric hydrogenation of the heterocyclic ring of a variety of isoquinolines was performed using a chiral iridium catalyst, giving a range of (*R*)-THIQ products in high yields (>80%) and high enantiomeric excesses (e.e.s) (90-99%). However, due to the resonance stability of the pyridine ring, hydrogenation must be performed under high pressure (*ca.* 80 bar), and the isoquinolines must be prepared synthetically (via Pictet-Gams or Bischler-Napieralski methods).[62,236] Iridium is also one of the lowest abundance elements on Earth, making its continued use unsustainable.[237] Oxidatively sensitive groups on the isoquinoline benzene ring are also not tolerated and so to access these compounds, a multi-step synthesis involving a protecting group strategy would have to be used.

Acid or phosphate-catalysed Pictet-Spengler reactions can also be employed to give racemic tetrahydroisoquinolines products, however regioselectivity can be poor. This means that labour-intense preparative-HPLC purification is often required to give single regioisomeric products.[8,15] The industrial production of the widely-prescribed pharmaceutical, solifenacin succinate (**81**) (Figure 5.1) involves the chiral resolution of racemic 1-phenyl-1,2,3,4-THIQ by tartaric acid recrystallisation. Although high enantiopurity in the product is achieved, the process is limited by poor atom economy and the THIQ scaffold must be generated via the Bischler-Napieralski reaction.[238] Isolating single enantiomer products of these compounds is meaningful because typically only one enantiomer possesses the desired medicinal properties.

Some enzymatic routes have been developed to generate 1-aryl THIQs, for example, using a stereoselective imine reductase (IRED).[156] Zhu *et al*. have reported the screening of a range of IREDs which are capable of generating a range of 1-aryl-THIQs in high yields and enantiopurities, with methoxy-, chloro- and methyl- substituents on the aryl ring. *Ortho* substituents were poorly tolerated, likely due to steric hinderance in the enzyme active site.[239] In a recent publication by Hestericová *et al*.,[156] an artificial metalloenzyme was also developed based upon the high affinity between biotin and streptavidin. Here, a chiral biotinylated iridium co-factor was irreversibly anchored into a streptavidin-binding protein. Directed evolution was used to develop two artificial transfer hydrogenases (ATHases) with opposing stereoselectivities. This system could generate both enantiomers of product in high enantiomeric excesses, cell-free extracts could be used, high substrate loading (100 mM) was tolerated, and the enzymes are stable under biphasic conditions. However, addition of an expensive diamide, 1,1-azobis(*N,N*-dimethylformamide) is required to prevent irreversible catalyst poisoning by glutathione, the imine starting material must be synthesised chemically and a relatively narrow substrate scope was presented. (1*S*)-1-Phenyl-THIQ (the precursor to the widely used anti-diuretic solifenacin succinate) was synthesised in 86% e.e.. Another enzymatic route to (1*S*)-1-phenyl-1,2,3,4-THIQ, developed by Ghislieri *et al*.[67] involves the use of an enantiospecific monoamine oxidase (MAO). The MAO used was engineered using directed evolution methods. Here, the racemic THIQ was generated chemically, and the MAO used to selectively oxidise a single enantiomer of the amine to give the imine. The imine was then reduced *in situ*, with ammonia borane to give the racemic amine. This results in the accumulation of the opposing single enantiomer of amine to the one that the MAO oxidises. A wide range of chiral amines were obtained in high yields (45 – 93%) and enantiomeric excesses (90 – 99%). Two disadvantages of the system are that firstly, oxidatively sensitive substrates (such as catechols) are not tolerated as reactions must be performed under an oxygen atmosphere. Also, the racemic substrate must first be prepared chemically, in a regioselective manner.

5.1.2 Naturally occurring 1-aryl-tetrahydroisoquinolines

Most examples of naturally occurring THIQs are benzylic at the C-1 position, generated in the NCS-mediated pathway to the benzylisoquinoline alkaloids. A variety of enantiopure 1-aryl-THIQs, known as the cryptostylines (**48 – 50**) were isolated from two species of orchid in the 1970s (as shown in Figure 5.2). Interestingly, in *Cryptostylis erythroglossa*, solely the (*R*)-enantiomer of product was isolated and in *Cryptostylis fulva.*, only the (*S*)-enantiomer was found (Figure 5.2).[114,115] The stereochemistry of the isolated products was determined by x-ray crystallography.[116]



(*S*)-(+)-Cryptostyline I **48** (*S*)-(+)-Cryptostyline II **49** (*S*)-(+)-Cryptostyline III **50**

Figure 5.2: The cryptostylines, a variety of naturally occurring enantiopure 1-phenyl-THIQs isolated from *Cryptostylis* fulva. The analogous *levo*rotatory compounds were isolated from *Crptostylis* erythroglossa.[114,116] Figure adapted from Roddan *et al.*[149]

Some feeding studies have been performed and determined that the amine required is dopamine, generated from tyrosine or phenylalanine in an analogous pathway to the biosynthetic routes in BIA-producing plants. The biosynthetic origin of the phenyl group at C-1 is unknown.[115] Due to the apparent isolation of single isomer products, it is likely that the biosynthesis involves an epimerase or Pictet-Spenglerase.

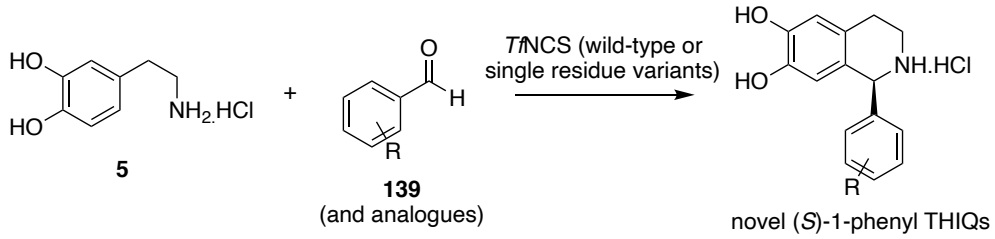
5.1.3 The use of norcoclaurine synthases to generate 1-aryl-tetrahydroisoquinolines

The work presented in Chapter 3, showed that despite previous reports, α-methyl substituted aldehydes are in fact well tolerated as substrates for *Tf*NCS. When a racemic aldehyde was used as the carbonyl substrate, a kinetic resolution of the

aldehyde was performed by the wild-type enzyme, giving THIQ products with (1*S*,1'*R*) stereochemistry in high diastereomeric ratios (up to 96:4). Two active site mutants, L76V and M97V were shown to improve diastereomeric ratios (up to 98:2) and conversions in products (up to >99%), respectively.[148] *Tf*NCS has also been shown to accept a variety of ketones as substrates, to generate 1,1-disubstituted THIQs. In particular, bulky substituted cyclohexanone derivatives were incorporated.[36] Therefore, it is not implausible that benzaldehydes may also be accepted as substrates.

Previous reports have shown that benzaldehyde derivatives were not well-tolerated as substrates by NCSs. Screening by Ruff *et al*. showed that only trace amounts of product are generated using benzaldehyde or 3-hydroxybenzaldehyde as substrates using wild-type *Tf*NCS.[21] Similar results were observed when using *Cj*NCS2 as the reaction catalyst.[22] In both cases, enzymatic assays were performed with low substrate loading (10 μM to 1 mM) and it is also possible that these could have been generated by non-enzymatic reactions. Subsequent studies have reported high K_d values of both substrates, indicating that high substrate loading is necessary for increased turnover of the substrates to give product.[26] Indeed, attempts to generate unnatural THIQ products using alpha-substituted amino acids (L-valine and L-isoleucine) *in vivo* using a yeast-based platform have been unsuccessful, likely due to the requirement of high substrate loading which is challenging to achieve in these systems.[130]

Therefore, it was hoped that using these active site mutants as a starting point, we could extend the known substrate scope of *Tf*NCS to also accept a variety of benzaldehydes as substrates and thus expand the portfolio of enantiopure THIQs that can be synthesised via an NCS-mediated reaction. This would also lead to the single-step syntheses of a range of novel (1*S*)-aryl-THIQs which are challenging to synthesise chemically due to the oxidatively sensitive catechol (Scheme 5.1).

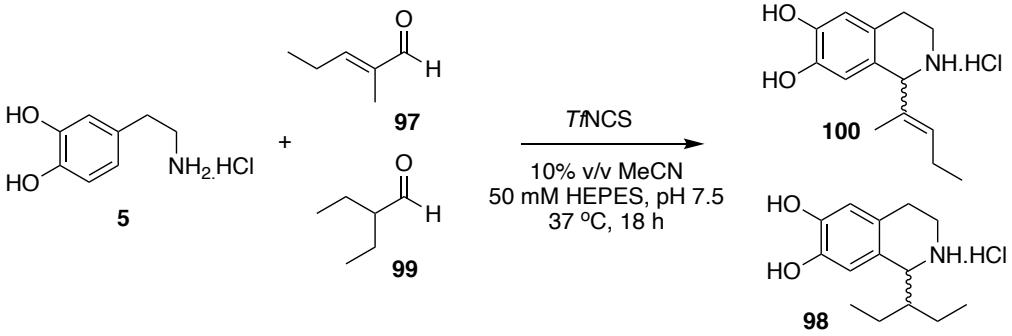


Scheme 5.1: **Summary of the project aim, to generate a variety of novel (1*S*)-aryl THIQ products in a single, regioselective and enantioselective step, using *Tf*NCS as a reaction catalyst.** This would be achieved from the selective Pictet-Spengler reaction between dopamine and a benzaldehyde analogue. Figure adapted from Roddan *et al.*[149]

5.2 Results and Discussion

5.2.1 Expanding the NCS substrate scope

Previous work (Chapters 3 and 4) has showed that α-methyl substituted aldehydes are indeed accepted as NCS substrates, despite previous reports that α-substituted aldehydes generally, are not accepted.[21,28] Two active site mutants of *Tf*NCS, L76V and M97V were utilised to improve diastereomeric excesses and conversions in products of reactions between dopamine and a variety of α-methyl substituted aldehydes. Based upon this, initial screens were performed using these promising mutants, compared with the wild-type enzyme with more challenging substrates, including 2-methyl-2-pentenal and an α-ethyl substituted aldehyde, 2-ethylbutanal, where low conversions were previously observed with the wild-type enzyme (Scheme 5.2).[148]



Scheme 5.2: NCS catalysed reactions performed between a conjugated α-methyl substituted aldehyde and an α-ethyl substituted aldehyde to give 1,1'-substituted THIQ products.

There is interest in accessing the THIQ product of the NCS reaction between dopamine and 2-ethyl-butanal as the racemic compound has been shown to act as a bronchodilator[163,240]. Conversions with these single-point mutants of *Tf*NCS (L76V, M97F, and M97V), compared with the wild-type enzyme are given in Figure 5.3.

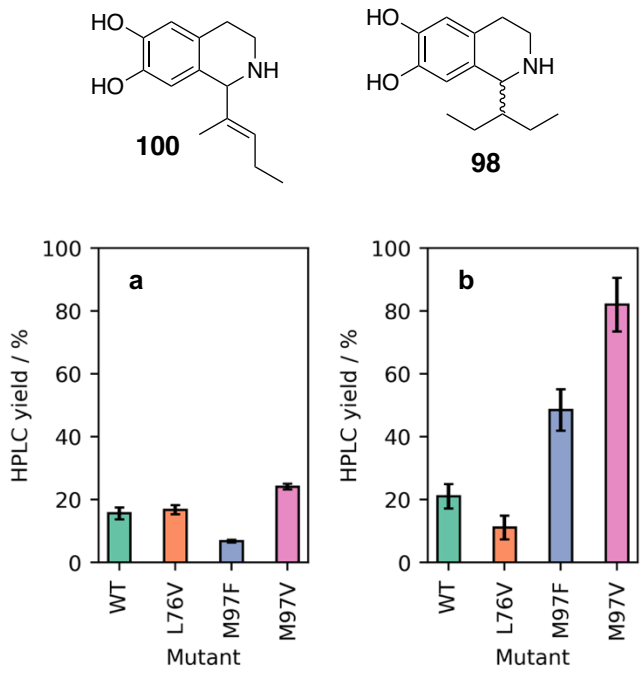


Figure 5.3: HPLC yields of reactions between dopamine and 97 or 99 using norcoclaurine synthases as the reaction catalyst. **a.** Reaction with 2-methyl-2-pentenal (**97**) to give **100. b.** Reaction with 2-ethylbutanal (**99**) to give **98**. *Reaction conditions*: dopamine HCl (10 mM), aldehyde (20 mM), sodium ascorbate (10 mM), purified *Tf*NCS (final concentration of 0.2 mg mL^{-1}, WT = Δ33*Tf*NCS, L76V = Δ29*Tf*NCS-L76V, M97F = Δ29*Tf*NCS-M97F and M97V = Δ29*Tf*NCS-M97V) in HEPES buffer (50 mM, pH 7.5) with MeCN (10% *v/v*) at 37 °C for 18 h. 200 μL scale reactions; samples were prepared by workup method 1, yields were determined by monitoring product formation against standards by analytical achiral HPLC (method 1). Figure adapted from Roddan *et al.*[149]

For enzymatic reactions between dopamine (**5**) and **100**, conversions were comparable when using the L76V mutant and the wild-type construct. Based on a co-crystallised structure of NCS gained with an α-methyl substituted reaction mimic (PDB: 6RP3), the 'R' group of the aldehyde sits close to this region in the active site, as shown in Figure 5.4.[148] However, due to the lack of rotational freedom of the aldehyde, it appears that this small increase in space is not sufficient to improve conversions with this substrate.

The methyl group of the aldehyde part of the mimic sits tightly in the small hydrophobic pocket between M97 and E110. This therefore rationalises why conversions are lowered with decreasing space in this region (M97F) and improved with increased space in this region (M97V), particularly for an aldehyde with restricted conformational freedom due to the alkene double bond. For conversions with 2-ethylbutanal (**99**) as the substrate, 50% lower conversions were observed with the L76V mutant, despite the creation of additional space in the active site. A similar observation was noted when the NCS mutant M97F was used as the reaction catalyst. This suggests that perhaps larger conformational changes occur in the enzyme upon the binding of bulkier aldehydes as substrates. As the crystal structure gained is simply a 'snapshot' of the mechanistic process, complete mechanistic insight cannot be gained from a single structure, although, it is known that the mechanism is highly dynamic from NMR-based studies.[26]

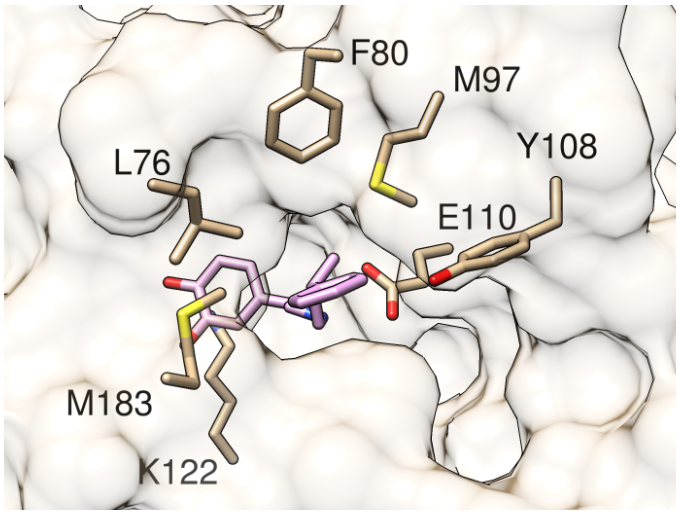
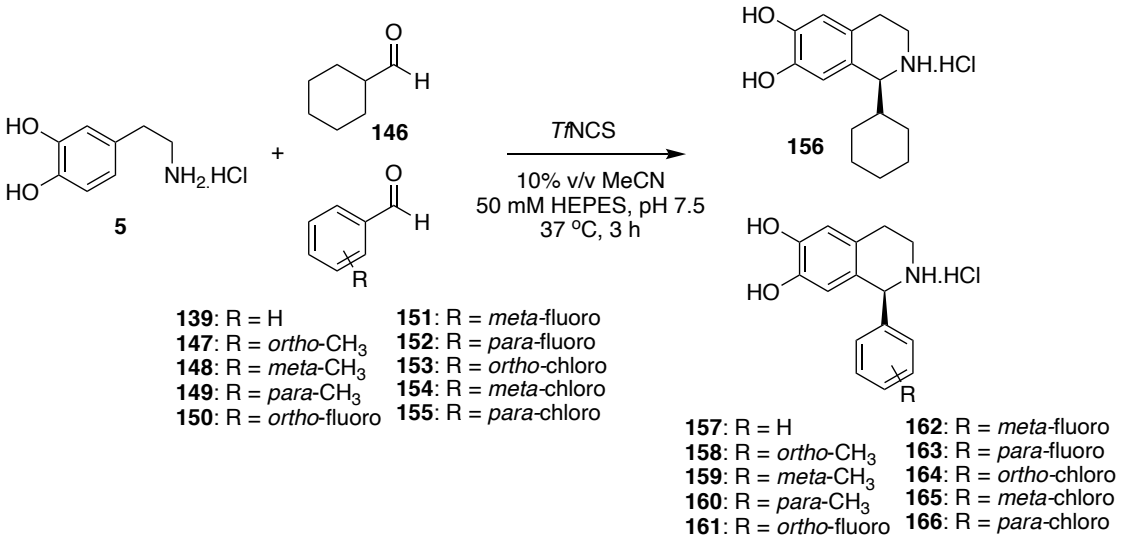


Figure 5.4: Co-crystallised structure of Δ33*Tf*NCS with alpha-methyl aldehyde reaction mimic bound in the active site. Figure was prepared using USCF Chimera using PDB: 6RP3.

Conversions were improved 4-fold with the single mutation, M97V, further highlighting the further applicability of this mutant with α-substituted aldehydes as substrates. Unfortunately, enantiomeric excesses of the products were not determined as these aldehydes were not converted to racemic products upon reaction with dopamine, using a phosphate-mediated, racemic Pictet-Spengler reaction.[15]

5.2.2 Acceptance of benzaldehyde and derivatives by NCS

In previous screens of reactions between dopamine and benzaldehyde (**139**), using *Tf*NCS as the reaction catalyst, performed by Ruff *et al*, products were not detected.[21] In these assays, substrate concentrations of 1 mM dopamine and 1 mM aldehyde were used, combined with very high enzyme concentrations (300 μM final concentration which is equivalent to 6.3 mg mL^{-1}). As demonstrated in previous reports, high substrate loading is well-tolerated by NCS and is indeed required to gain high conversions of the product. NMR titration experiments were performed by Berkner *et al*. to further elucidate the NCS reaction mechanism. The K_d value of dopamine was found to be 5 mM, thus indicating very weak substrate binding.[26] With this in mind, assays performed in this study used increased substrate loading; 10 mM amine and 20 mM aldehyde (**139, 146-155**). High concentrations of aldehyde were used (2 e.q.), with dopamine as the limiting reagent. An excess of aldehyde was used in enzymatic reactions because the aldehyde degrades with time and so use as the limiting reagent would give inaccurate quantitative conversions.



Scheme 5.3: A variety of benzaldehyde analogues were tested as substrates for *Tf*NCS using dopamine as the amine substrate. *Reaction conditions*: dopamine·HCl (10 mM), aldehyde (20 mM), sodium ascorbate (10 mM), purified Δ29*Tf*NCS-M97V (0.2 mg mL^{-1} final concentration) in HEPES buffer (50 mM, pH 7.5) with MeCN (10% v/v) at 37 °C for 3 h, 100 μL

scale reactions; samples were prepared by workup method 1, yields were determined by monitoring product formation against standards by analytical achiral HPLC (method 1).

Assays were performed using benzaldehyde (**139**) and a variety of analogues (**146-155**) using **5** as the amine substrate (Scheme 5.3). Conversions to the products (**157-166**) using wild-type enzyme and the mutants L76V, M97F and M97V are given in Figure 5.5**.** Reactions were performed for 3 hours to minimise the contribution of the spontaneous background Pictet-Spengler reaction between the two substrates, as the aldehyde is highly activated. The major enantiomer of product formed was assumed to have *S* stereochemistry, based upon previous reports, and was later confirmed by derivitisation of **157** to a known standard. The background reaction, as observed in the no enzyme controls, was shown to give racemic THIQ products by chiral HPLC.

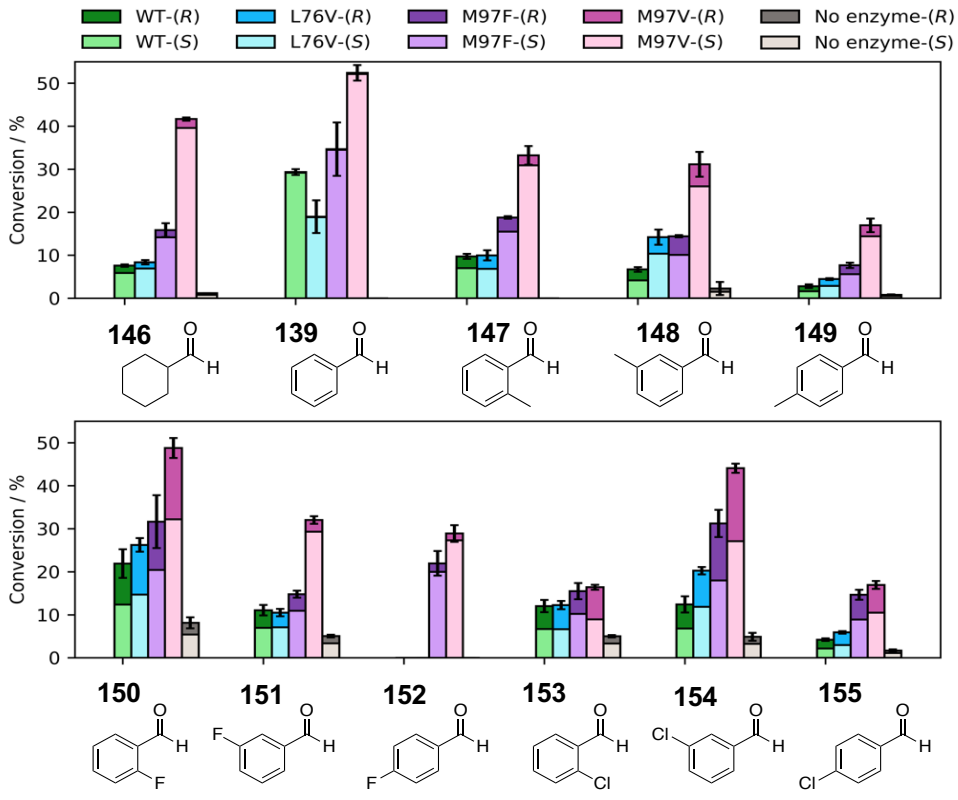


Figure 5.5: Conversions of reactions between dopamine and a variety of benzaldehyde analogues. The aldehyde substrates used (**139, 146-155**, Scheme 5.3) are given on the x axis. Wild-type or active site mutants of NCS were used as the reaction catalyst with reaction conditions given in Scheme 5.3. Enantiomeric excesses in the products are given by showing the amounts of (*R*)- and (*S*)- product generated through the reactions, determined by chiral HPLC analysis (method 4 (for **156** and **157**) and 5 for all other compounds). Conversions of the reactions were determined by analytical achiral HPLC analysis based upon product formation,

compared to calibration curves of purified standards (method 1), see Appendix. Figure adapted from Roddan *et al.*[149]

Comparing the difference in conversions of reactions between the aromatic aldehyde, benzaldehyde (**139**) and the aliphatic analogue, cyclohexanecarboxyaldehyde (**146**), increased conversions were observed with **146** for all NCSs tested. Likely little to no background reaction occurs with **146** as the aldehyde substrate due to less activation of the aldehyde, by donation from the neighbouring pi-system, compared with the aromatic aldehyde, **139**. Similar results are observed when other aliphatic aldehydes are used as substrates. High enantiomeric excesses in the products are therefore observed with **139** as the aldehyde substrate (>99% for all *Tf*NCS variants).

In all cases, conversions with the wild-type enzyme and L76V with aldehydes **139, 146-155** are comparable and conversions with both substitutions at residue 97 resulted in improved yields. This is particularly interesting with the M97F variant as due to the bulky phenyl group, there is less space in this region of the active site.[148] In addition, in previous studies with ketones used as NCS substrates, in all cases, mutations at the M97 position (to phenylalanine, leucine or valine) resulted in poorer yields compared with the wild-type enzyme as the catalyst.[36] With the M97V variant, in a majority of cases, lower conversions were gained than with M97F. Previous NMR studies of the wild-type *Tf*NCS enzyme, have indicated it to be highly dynamic during the catalytic process with substrate binding resulting in large conformational changes across the entire protein.[26] The increased conversions and enantiomeric excesses in products in the M97F and M97V could therefore be due to an effect whereby when faster binding of the substrates occurs (by increasing hydrophobicity or active-site space, respectively) and turnover is faster, so the background Pictet-Spengler reaction is less efficient and so product enantiopurity is improved.

The effects of electron-donating (**147-149**) and electron-withdrawing substituents (**150-155**) on the aldehyde aromatic ring were tested by performing reactions with ortho-, *meta*- and *para*- substituted benzaldehyde analogues. It was hoped that altering the substitution on the phenyl ring at different positions could help to identify how the various enzyme mutants were affecting conversions. For reactions with all substituted benzaldehydes (**147-155**), enantiomeric excesses in the products were decreased compared with benzaldehyde (**139**), despite often low or no observed background reactions. This suggested that the lower stereoselectivities were associated with the presence of the enzyme, rather than any background reaction. It could be questioned whether this is due to reactivity in the enzyme active site, whereby formation of a small amount of the (*R*)-enantiomer of product is formed or perhaps there is an alternative hydrophobic surface on which the racemic reaction can occur. Further elucidation of this would be challenging and there are also issues associated with quenching both the enzymatic and background reactions, as discussed later.

Conversions for all *ortho*- (**147, 150** and **153)** and *meta*- (**148**, **151**, and **154**) substituted benzaldehydes tested were comparable to those with benzaldehyde, whereas *para*-substitutions (**149**, **152** and **155**) were less well tolerated. Interesting with *para*-fluorobenzaldehyde (**4h**) as a substrate, no conversions were observed with the wild-type enzyme but with the two M97 mutants tested, enzymatic activity was observed. Particularly poor enantiomeric excesses were observed for reactions with all chlorinated benzaldehyde derivatives (**153 - 155**).

To understand which features of the reaction conditions were required for successful conversions, a reaction towards **157** was performed using the conditions described by Ruff *et al.*[21] The reaction conditions were then adjusted towards the conditions described earlier in Figure 5.5 by increasing substrate concentrations, enzyme concentrations and reaction temperatures. Resulting product yields with the various

reaction conditions are given in Figure 5.6. Product yields are given in percentages of product formation, compared with the maximum theoretical percentage yield based upon calibration curves of isolated **157.** Reaction temperature has little effect on yield and a minor increase in yield is observed with increased substrate loading. 3.8 μM is approximately 0.08 mg mL^{-1} of *Tf*NCS, so half the amount of enzyme was used compared with the conditions described in Figure 5.5. However, a 10-fold increase in enzyme concentration in the conditions used by O'Connor does not improve the yields to that observed in this work (11% previously vs. 30% here). Benzaldehyde is highly oxidatively sensitive and only one equivalent was used previously.[21] To accommodate for this oxidation and ensure that at least one equivalent of aldehyde was present compared to phenethylamine, in this work, two equivalents of aldehyde were used which may account for the increased yields observed. A different co-solvent was also used here (acetonitrile vs. methanol previously) which also may have an effect. Increased conversions are observed when using DMSO as a co-solvent when ketones are used as substrates, compared with using acetonitrile.[36]

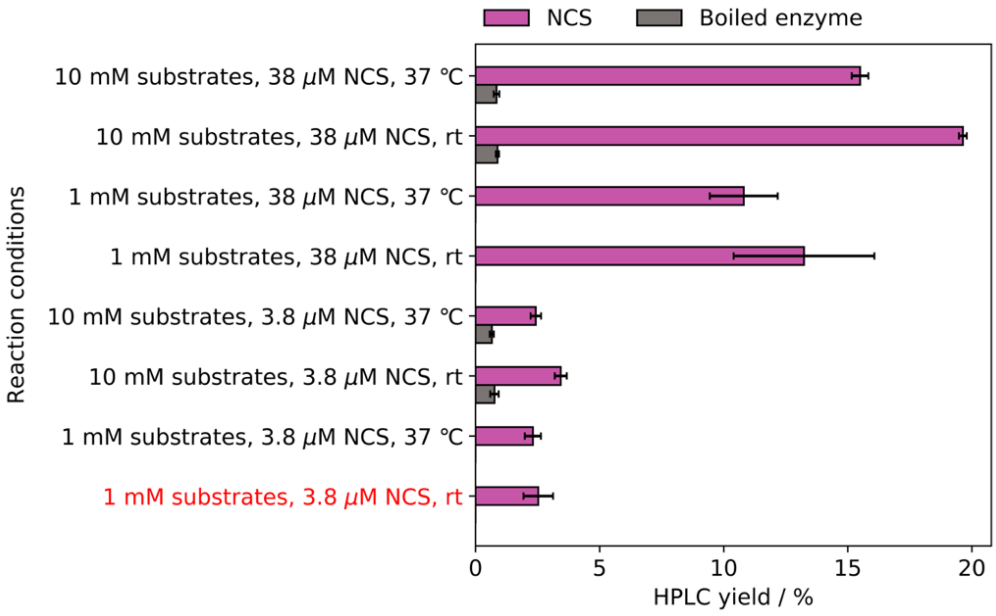


Figure 5.6. Yields of product formation of 157 following reaction conditions described by O'Connor *et al.*[21] The reaction conditions performed by O'Connor *et al.* are given in red. Reaction conditions were also altered to determine the cause of low conversions. *Reaction conditions*: Stock solutions of dopamine (100 mM in MeOH) and benzaldehyde (100 mM in MeOH) were prepared and added to Tris buffer (100 mM, pH 7) and Δ29*Tf*NCS (stock at 480

μM in 20 mM Tris, 50 mM NaCl, pH 7.5). Reactions were prepared to 100 μL and were performed for 3 h. Boiled enzyme was used as the negative control. Reactions were quenched by addition of 100 μL MeOH, centrifuged (13,000 rpm, 10 min, 4 °C) and the supernatant analysed by RP-HPLC (method 1). Conversions were determined based upon a calibration curve of **5b** (see Appendix). Reactions were performed in triplicate and standard deviations reported. Figure reproduced from Roddan *et al.*[149]

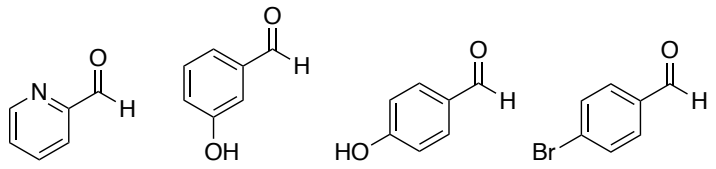
These initial attempts to perform assays with the benzaldehydes highlighted the issues of the spontaneous, racemic background reaction occurring and thus the method of quenching the enzymatic reaction needed to be considered. Purified enzyme had been used to eliminate any phosphates from the cell-lysate which would catalyse the background reaction. Previous enzymatic assays performed using *Tf*NCS have always used the addition of acid (aq. HCl or TFA) to precipitate the enzyme.[21,36,118] However, acidic conditions are also widely known to catalyse the Pictet-Spengler reaction.[8] This has not previously been an issue with NCS reactions the as previously used carbonyl substrates are not particularly activated and mildly acidic conditions at low temperatures are unlikely to catalyse the racemic background reaction and thus diminish the enantiopurity of the isolated products.

Due to the highly reactive nature of the benzaldehydes, alternative methods to quench the reactions were considered. Other commonly used methods to precipitate proteins include the addition of base or organic solvents. Dopamine (and analogues) are known to undergo radical polymerisation reactions under basic conditions so observed conversions would be lower than expected.[241] DMSO could not be used as it is a known oxidising agent and the high volatility of methanol would affect the quantitative analysis of conversions. Therefore, an equal volume of acetonitrile to reaction volume was used to quench enzymatic reactions. Due to these considerations, acetonitrile was also used as the co-solvent for enzymatic reactions rather than DMSO as previously reported. Previous studies using α-methyl substituted aldehydes as substrates showed that comparable conversions are observed with either co-solvent (Figure 3.1).

After precipitation of the enzyme, it was not possible to stop the spontaneous background reaction. Therefore, to reduce this reaction, samples were frozen (-20 °C) after quenching with acetonitrile and analytical RP-HPLC analysis was performed with samples maintained at 4 °C. To halt the reactions and prepare samples for chiral HPLC analysis, samples were saturated with sodium bicarbonate, to adjust to pH to 8 and immediately extracted into ethyl acetate. The pK_a of dopamine amine is around 9, so the majority would remain in the aqueous phase and only product and benzaldehyde extracted into ethyl acetate, meaning less amine available to participate in a background Pictet-Spengler reaction. This method however is not appropriate for determining conversions to the product as some may be retained in the aqueous phase (and has been shown by RP-HPLC analysis of the aqueous phase after extraction) and therefore lost during the extraction process.

5.2.3 Attempts to extend the benzaldehyde substrate scope of NCS

With a view to further extending the benzaldehyde substrate scope beyond simple methylated or halogenated analogues, a variety of bulkier, activated benzaldehyde derivatives (given in Figure 5.7) were tested as substrates for *Tf*NCS.



167 **168** **169** **170**

Figure 5.7: **A range of other benzaldehyde analogues tested as substrates for *Tf*NCS using dopamine as the amine substrate**. *Reaction conditions*: dopamine·HCl (10 mM), aldehyde (20 mM), sodium ascorbate (10 mM), purified Δ29*Tf*NCS-M97V (0.2 mg mL^{-1} final concentration) in HEPES buffer (50 mM, pH 7.5) with MeCN (10% v/v) at 37 °C for 3 h, 100 µL scale reactions; samples were prepared by workup method 1, product formation was determined by analytical achiral HPLC (method 1).

For enzymatic reactions using *Tf*NCS (M97V) between dopamine and 2-pyridinecarboxaldehyde (**167**), product was observed by LC-MS, however a high background reaction was also observed in the control reaction and additional, unassignable peaks were also seen by aHPLC. The presence of the pyridine ring is likely causing side reactions due to activation of the aldehyde carbonyl by the electron withdrawing nitrogen.

Trace conversions were observed for reactions between dopamine and the two hydroxylated benzaldehydes (**168** and **169**), comparable to conversions observed in the control samples. *Ortho* and *para* isomers of the product observed by RP-HPLC were in the same ratios as observed for the control reaction, with higher peak areas in the enzymatic reaction. Trace conversions were also observed with the brominated aldehyde (**170**). It is likely that these larger hydroxyl- (**168** and **169**) and bromo-substituted aldehydes (**170**) are not accepted as intermediates simply do not fit into the active site of the enzyme.

5.2.4 Improvement of enantiomeric excesses of the products

Although impressive conversions were observed for reactions using the M97V mutant, for some substrates, the enantiomeric excesses observed in the products were lower than expected due to the high racemic background reaction occurring simultaneously. It was hoped that by altering the concentrations of enzyme used and by stopping the reaction at different time points, optimal reaction conditions could be found to achieve both high conversions and high enantiomeric excesses in the products without using large quantities of enzyme. Timepoint assays were performed for the reactions between dopamine and benzaldehyde (**139**) (Figure 5.8) using varying concentrations of Δ29*Tf*NCS-M97V as the reaction catalyst.

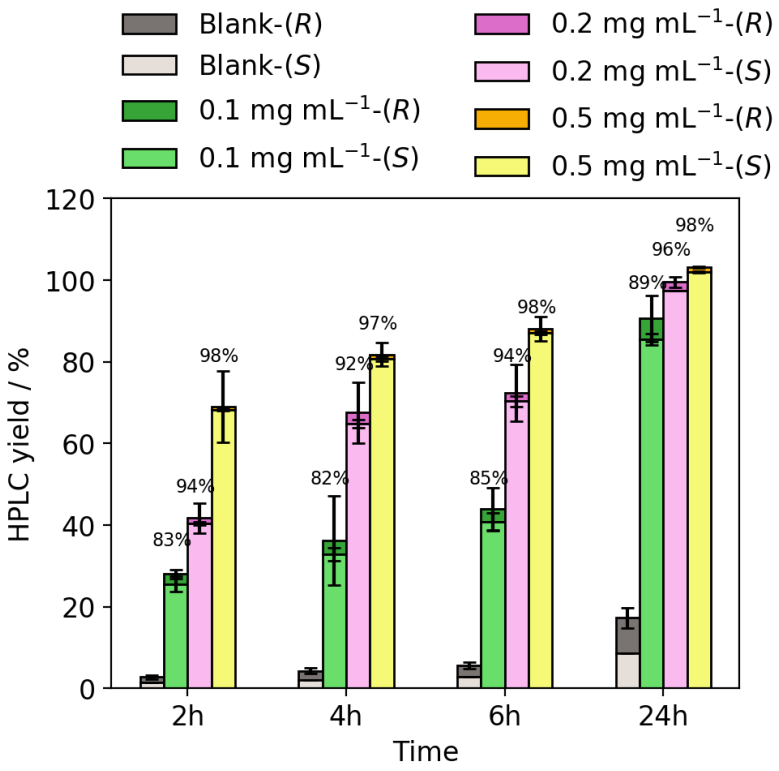


Figure 5.8: **Conversions and product enantiopurities for reactions between dopamine and benzaldehyde (139) to give 157, catalysed by *Tf*NCS-M97V.** Varying timepoints and final concentrations of enzyme were used. *Reaction conditions*: 10 mM dopamine HCl, 10 mM sodium ascorbate, 20 mM benzaldehyde in HEPES buffer (50 mM, pH 7.5) with 10% *v/v* MeCN and final concentration of Δ29*Tf*NCS-M97V *Tf*NCS at 0.1, 0.2 or 0.5 mg mL^{-1}. Control reactions were performed using the same conditions but the *Tf*NCS sample was substituted for enzyme buffer (20 mM Tris, 50 mM NaCl, pH 7.5). Reactions were performed at 37 °C for 3 h, 100 µL scale reactions; samples were prepared by workup method 1, yields were determined by monitoring product formation against standards by analytical achiral HPLC (method 1). Above each bar, the enantiomeric excess of **157** formed is given as a percentage. Figure adapted from Roddan *et al.*[149]

As shown in Figure 5.7, the background reaction occurs between the two substrates at all timepoints. Indeed, in the absence of enzyme, a 17% conversion was observed after 24 h. High enantiomeric excesses (97-98%) in the products were observed for all reactions using 0.5 mg mL^{-1} (final concentration) of enzyme. However, using 0.2 mg mL^{-1} enzyme gave 95% e.e. in the product at both 6 h and 24 h. Therefore, since conversions were comparable at 24 h, a 6 h reaction with a 0.2 mg mL^{-1} final concentration of enzyme would minimise the enzyme concentration required, reducing

the protein preparation requirements. However, purification of the reaction sample prepared using a reaction time of 24 h and a final enzyme concentration of 0.5 mg mL^{-1} would be simplest as all dopamine was consumed by that stage. This would mean that purification by an extractive workup may be possible.[118] A 24 h reaction time was therefore selected for performing these reactions with other benzaldehyde analogues.

To optimise the reactions between dopamine and a substituted benzaldehyde, which had proved more challenging for gaining high enantiomeric excesses in the products (see Figure 5.5), a similar timepoint assay was performed for the reaction between dopamine (**5**) and 3-methylbenzaldehyde (**148**) (Figure 5.9). For an NCS-mediated reaction with an enzyme concentration of 0.1 mg mL^{-1} and a reaction time of 3 h (Figure 5.5), the product was obtained in poor enantiopurity (67% e.e.) and low yield (31%). Therefore, reactions were focussed on using a 0.2 or 0.5 mg mL^{-1} final concentration of enzyme. Again, a high racemic background reaction was observed (15% conversion after 24 h). However, the amount of (*R*)-product formed was higher than would be expected based upon the no enzyme reactions, suggesting that a racemic Pictet-Spengler reaction is associated with the use of enzyme, but high concentrations of enzyme are also required for fast turnover of the substrates to give high e.e.s. For all reactions with 0.2 mg mL^{-1} of enzyme e.e.s of 74-78% were observed, whereas with 0.5 mg mL^{-1} enzyme, e.e.s of 84-85% were observed at 2-6 h, increasing to an e.e. of 87% at 24 h, combined with very high conversion levels. It was therefore likely that using these reaction conditions with the other substituted benzaldehyde analogues enantiopurities might be improved. Clearly, higher concentrations of pure enzyme (0.5 mg mL^{-1}) are optimal to achieve high enantiomeric excesses in the (1*S*)-1-aryl-THIQs generated by an NCS-catalysed reaction.

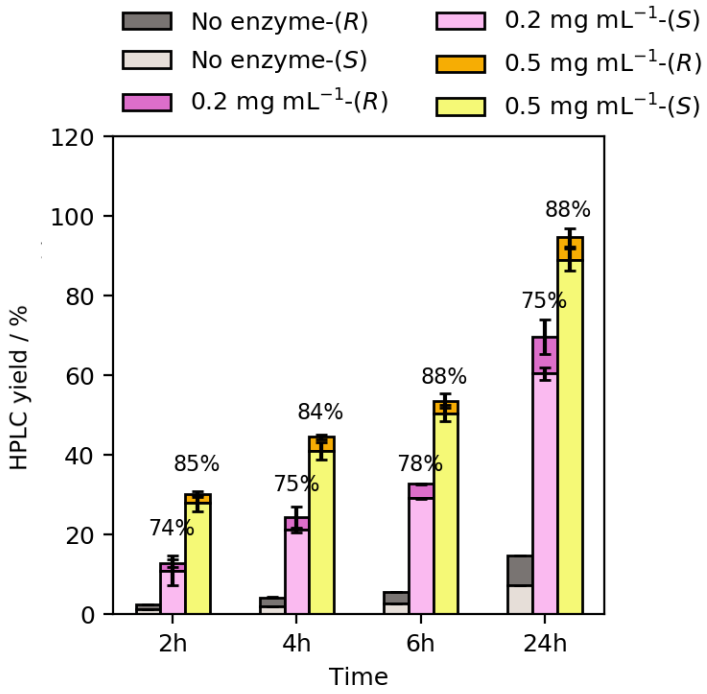


Figure 5.9: **Conversions and product stereochemistry for reactions between dopamine (5) and 3-methylbenzaldehyde** (**148**) **to give 159**. *Tf*NCS-M97V at varying final concentrations was used as the reaction catalysts. *Reaction conditions*: 10 mM dopamine HCl, 10 mM sodium ascorbate, 20 mM benzaldehyde in HEPES buffer (50 mM, pH 7.5) with 10% *v/v* MeCN and final concentration of *Tf*NCS at 0.2 or 0.5 mg mL^{-1}. Control reactions were performed using the same conditions but the *Tf*NCS sample was substituted for enzyme buffer (20 mM Tris, 50 mM NaCl, pH 7.5). Reactions were performed at 37 °C for 3 h, 100 µL scale reactions; samples were prepared by workup method 1, yields were determined by monitoring product formation against standards by analytical achiral HPLC (method 1). Enantiomeric excesses of **159** formed are given above each bar as a percentage. Figure adapted from Roddan *et al.*[149]

5.2.5 Optimised NCS-catalysed reactions with a range of benzaldehydes as substrates

Reactions performed at the higher enzyme concentrations (0.5 mg mL^{-1}) gave high enantiopurities in the products of reactions between dopamine (**5**) and benzaldehyde (**139**) or 3-methylbenzaldehyde (**148**) (Figures 5.8 and 5.9). Therefore, to improve the enantiomeric excesses the products with other benzaldehyde analogues, NCS-mediated reactions were performed using the most promising mutant, Δ29*Tf*NCS-M97V at a final enzyme concentration of 0.5 mg mL^{-1} for 24 h. Conversions with the amounts of (*R*) and (*S*) products generated for each are given in Figure 5.10.

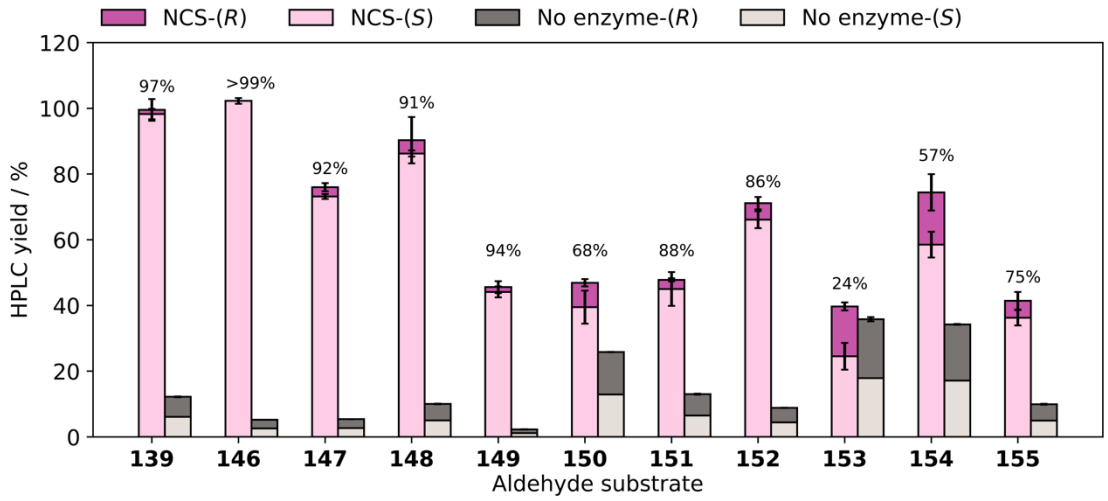


Figure 5.10: **Yields and enantiomeric excesses of products generated by an Δ29*Tf*NCS-M97V catalysed reaction between dopamine and a variety of benzaldehyde derivatives.** *Reaction conditions*: dopamine·HCl (10 mM), aldehyde (20 mM), sodium ascorbate (10 mM), purified Δ29*Tf*NCS-M97V (0.5 mg mL^{-1} final concentration) in HEPES buffer (50 mM, pH 7.5) with MeCN (10% v/v) at 37 °C for 24 h, 100 µL scale reactions; control reactions were performed using the same conditions but the *Tf*NCS sample was substituted for enzyme buffer (20 mM Tris, 50 mM NaCl, pH 7.5). Yields were determined by monitoring product formation against standards by analytical RP-HPLC (method 1, see Appendix). Enantiomeric excesses of the THIQs formed are given above each bar as a percentage.

As previously shown, complete conversions were observed for reactions with **139** and **146** to give **156** and **157**, with both products formed in excellent enantiopurity (97% and >99%, respectively). Using these conditions also significantly improved both the conversions and enantiopurity of the products with methylated benzaldehyde derivatives (**147-149**) where conversions of 76% and 90% were observed for the *ortho*- (**158**) and *meta*- (**159**) substituted compounds, respectively in >90% e.e. Lower enantiopurity and conversions were observed with the *para*-substituted analogue (**160**), (68% e.e. and 45% yield). Despite the observation of a high racemic background reaction for NCS-catalysed reactions involving chlorinated benzaldehyde derivatives (**153** and **153**), reasonable enantiomeric excesses in the products were obtained (68-88%). An increasingly significant background reaction was observed for reactions with chlorinated benzaldehydes. Interestingly, a similar trend was observed with the halogenated analogues; lower background reactions were observed with the *para* <

meta < *ortho*. Presumably, this is due to a combination of sterics and carbonyl activation. Despite this, the *para*-chlorinated product, **155** was generated in reasonable yield, 41% compared with only 18% with a reaction time of 3 h and enzyme concentration of 0.2 mg mL^{-1}. Increased enantiopurity in the product was also observed, 75% vs. 21% (Figure 5.9 compared with Figure 5.4).

Gaining high enantiomeric excesses in the products for reactions with halogenated benzaldehydes (**5f-k**) still proved challenging. It was hypothesised that this was due to a combination of a high background reaction (as the aldehydes are highly activated) and increased lability of the C-1 proton, due to the presence of electron-withdrawing halogen atoms on the aryl ring. To minimise racemisation at the C-1 position, reactions were performed using analogous conditions to those performed for Figure 5.10, except a lower pH (6 vs. 7.5) was used. Conversions and enantiomeric excesses of the resulting products (**161-166**) from reactions with aldehydes **150 – 155**, respectively) are given in Figure 5.11.

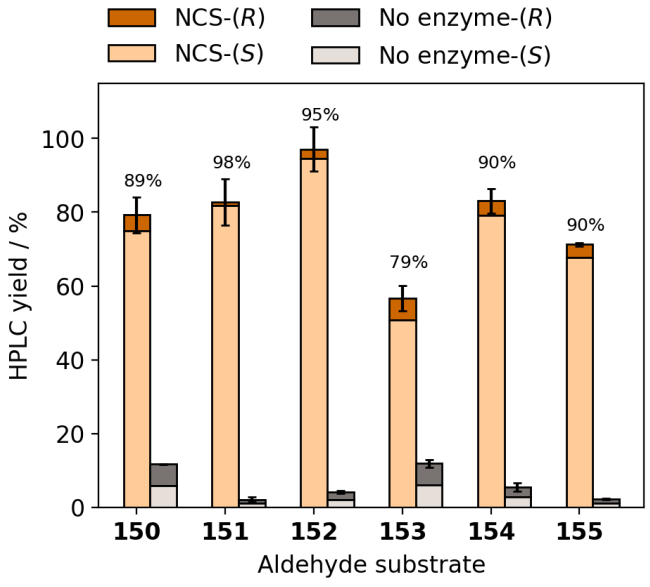


Figure 5.11: Yields and enantiomeric excesses of products generated by an Δ29*Tf*NCS-M97V catalysed reaction between dopamine and halogenated benzaldehyde derivatives at pH 6. *Reaction conditions*: **5** (10 mM), aldehyde (20 mM), sodium ascorbate (10 mM), purified Δ29*Tf*NCS-M97V (0.5 mg mL^{-1} final concentration) in HEPES buffer (50 mM, pH 6) with MeCN (10% v/v) at 37 °C for 24 h, 100 µL scale reactions; control reactions were performed using the

same conditions but the *Tf*NCS sample was substituted for enzyme buffer (20 mM Tris, 50 mM NaCl, pH 7.5). Yields were determined by monitoring product formation against standards by analytical RP-HPLC (method 1). Enantiomeric excesses of the THIQs formed are given above each bar as a percentage.

Remarkably, in all cases, improved yields and enantiomeric excesses in the products were observed. With the *ortho*-chlorinated product (**164**), at pH 7.5, the yield and e.e. observed were 40% and 24%, respectively, whereas at a lower pH, both were improved to 57% and 79% respectively. Interestingly, it appears that this was due to the minimal background reaction observed, which was unexpected, considering that the Pictet-Spengler reaction is catalysed by acidic conditions. Previously, reactions catalysed by *Tf*NCS have been performed at near neutral pH (7-7.5) and it is unclear how performing reactions at pH 6 improves both the enantiopurity and conversions with these substrates. Perhaps increased activation of the aldehyde helps improve substrate turnover or this NCS variant (M97V) is simply more active at a lower pH. Since the background reaction is minimized at pH 6, probably due to protonation of the HEPEs buffer, which may help activate the substrates. giving racemic Pictet-Spengler products, particularly with the activated benzaldehydes. This therefore means that more substrate is available for NCS reactions, leading to increased *S* product formation. It is therefore clear that reaction conditions need to be optimised for each substrate type, also considering that decreasing pH to 6 did not improve conversions with *N*-methyl-phenethylamines as substrates, as discussed in Chapter 6.

5.2.6 Synthesis of racemic 1-aryl tetrahydroisoquinolines

To generate racemic standards of the 1-aryl-THIQs for chiral HPLC analysis, the phosphate-mediated Pictet-Spengler reaction between dopamine and a variety of benzaldehyde derivatives (**139, 146** – **155**) was investigated[15]. Reactions were performed using previously reported conditions. Conversions were lower than expected (14 - 86%), as shown in Figure 5.12, considering the apparent spontaneous reaction between the two substrates without a reaction catalyst that was observed during

enzymatic reaction controls. Significant precipitation was observed in the reaction samples and additional peaks were also observed during RP-HPLC analysis that did not correspond to the THIQ products isolated from enzymatic reactions. This suggests that side reactions (e.g. oxidations) occur due to the highly reactivity of the benzaldehydes. To minimise this, reactions were performed under argon, and it is likely that some of the side-products are adhering to the plastic Eppendorf tubes that the reactions were performed in. In future attempts, these reactions could be performed in glass reaction vials instead.

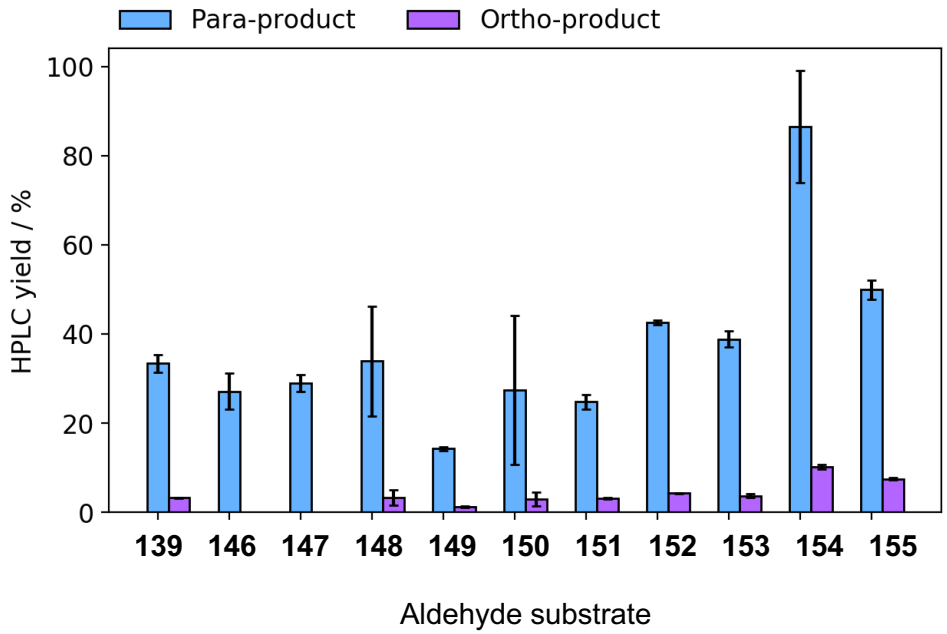


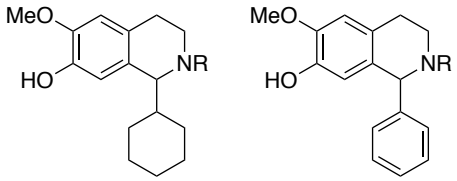
Figure 5.12: Conversions of phosphate-mediated reactions between dopamine and a variety of benzaldehyde analogues. *Reaction conditions*: 10 mM dopamine HCl, 10 mM sodium ascorbate, 20 mM aldehyde in KPi buffer (300 mM, pH 6) with 50% *v/v* MeCN. Reactions were performed on a 200 μL scale for 18 h at 60 °C. Samples were prepared by workup method 1. Yields were determined by monitoring product formation against standards by analytical RP-HPLC (method 1, see Appendix). Figure adapted from Roddan *et al.*[149]

A phosphate-mediated reaction is an excellent way of generating racemic THIQs under mild conditions, however one limitation is the lack of regioselective control with some aldehydes as substrates. For reactions between dopamine and linear or benzylic aldehydes, a ratio of the *ortho*:*para* isomers of product of 1:3 was observed, so single regioisomer products are typically isolated by preparative-HPLC. For reactions between

dopamine and α-methyl substituted aldehydes, only the *para* regioisomer was generated, likely due to steric hinderance. It was therefore hoped that using benzaldehydes, with increased bulk at the alpha-position, single regioisomer products would be formed. The conversions and regioisomer ratios of phosphate mediated reactions between dopamine and a variety of benzaldehyde analogues are given in Figure 5.12. For the majority of substrates, a ratio of the two regioisomers, *ortho:para* of 8-10:1 was observed.

5.2.7 Increasing molecular complexity using regioselective catechol O-methyltransferases

A variety of *meta*-methylated, *N*-acylated 1,2,3,4-THIQs (Figure 5.13) have been shown to down-regulate the production of various inflammatory cytokines and other toxic molecules in activated microglial cells. This activation is associated with a variety of neurodegenerative diseases such as Huntington's and Parkinson's disease. Several of these THIQs are under patent for preventing or treating a variety of degenerative or inflammatory diseases.[242] Performing the regioselective *meta*-methylation of a catechol is challenging chemically, as there is a lack of regioselective control and synthetic routes therefore require the use of a precursor with the correct methylation pattern and/or employ a multi-step, protecting group strategy. However, in the production of THIQs using NCS, the *meta*-hydroxyl group of the dopamine catechol is currently required for a productive enzymatic reaction, so the regioselective methylation reaction must be performed after the Pictet-Spengler reaction has occurred if meta-methylation is required.[21,28] If *para*-methylation is required, *para*-methylated dopamine could be used but it is not currently commercially available and cannot be generated by a catechol *O*-methyltransferase in complete conversion, so isolation of the pure amine would be challenging.[209]



R= Me, Et, Pr, Bu, CH_2Ph, $COCH_3Ac$

Figure 5.13: Examples of some compounds under patent for the treatment of a range of neurodegenerative diseases. Single enantiomer analogues of these compounds may be generated by an NCS-mediated reaction followed by methylation with a regioselective *O*-methyltransferase.

Two regioselective catechol *O*-methyltransferases, *Rn*COMT and *Mx*SafC have already been explored with the THIQ product of reaction between dopamine (**5**) and an α-methyl substituted aldehyde, as discussed in Chapter 4. Both enzymes were regioselective towards the C6/*meta*-position. Another *O*-methyltransferase, from *Thalictrum flavum,* 6OMT was also available and was being explored by other group members (Fabiana Subrizi and Benjamin Thair, Hailes group) for benzylic THIQs as substrates. 6OMT ((*S*)-norcoclaurine 6-*O*-methyltransferase), is involved in the same biosynthetic pathway as NCS towards the *in planta* production of BIAs.[197] 6OMT selectively methylates the C6-hydroxyl of (*S*)-norcoclaurine (**7**) to give (*S*)-cocolaurine (**8**) (Scheme 1.1). The substrate scope of 6OMT has not been widely explored however (*S*)-norlaudanosauline (**15**) is also accepted and there have been crystallographic investigations. In the co-crystallised structure gained of 6OMT with active-site bound SAH and (*S*)-norlaudanosauline (**15**), a hydrogen-bonding interaction can be observed between the benzylic hydroxyl groups and Asp-169.[243] It is unknown whether this interaction is essential for a productive reaction.

However, it was hoped that here, that investigation of three different, appropriate *O*-methyltransferases, would lead to the selective C6-hydroxyl methylation of (*S*)-1-aryl-THIQs towards the pharmaceutically relevant molecules shown in Figure 5.12. *Rn*COMT and *Mx*SafC were prepared as discussed in Chapter 4. 6OMT was expressed recombinantly in the same way at the other two enzymes, except after induction with

IPTG, the temperature was lowered to 16 °C and the cultures left to express for 48 h, as expression trials (performed by Fabiana Subrizi) showed that this increased yields. However, as demonstrated in Figure 5.14, poor expression was still achieved.

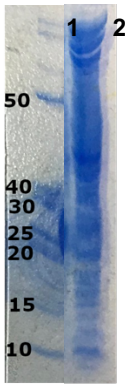
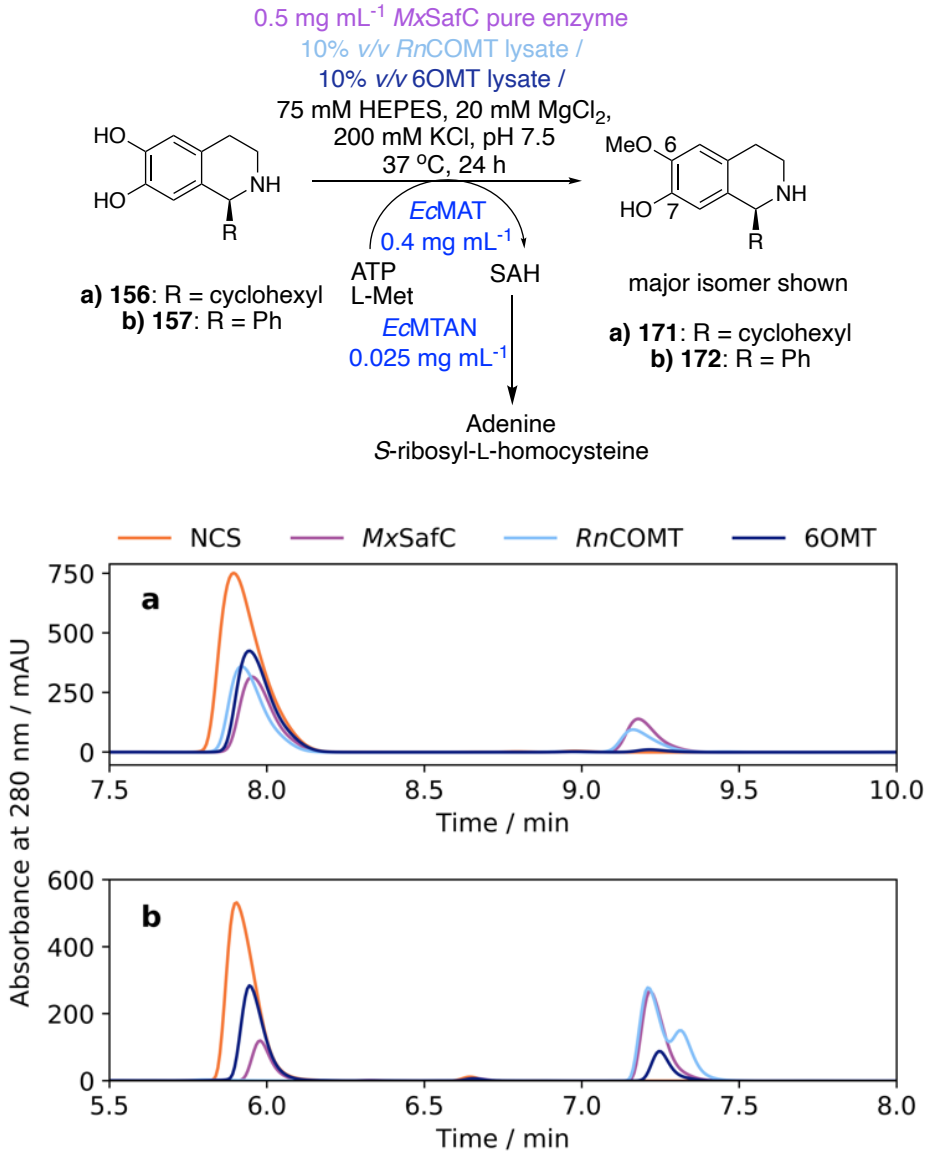


Figure 5.14: SDS-PAGE analysis of 6OMT expression. Lanes: 1, Benchmark™ protein ladder (masses given in kDa), 2. Clarified cell lysate of 6OMT expression. Protein expected at 39 kDa.

The three *O*-methyltransferases (*Rn*COMT, *Mx*SafC and 6OMT) were tested with two (*S*)-1-substituted THIQ products (**5a** and **5b**) generated from the NCS reaction. Achiral HPLC traces of the reactions performed are shown in Scheme 5.4. The analytical HPLC trace of the product of the NCS reaction is also shown for comparison, with the methylated products eluting over a minute later, as expected due to the decrease in polarity.



Scheme 5.4: **Achiral HPLC traces of regioselective C6-*O*-methylation of two THIQ products. a. 171, b. 172.** *Reaction conditions:* Once the NCS reaction was completed (performed on 100 μL scale), the reaction mixture was quenched and lyophilised (workup method 5). The residue prepared in 25 mM HEPES, 20 mM $MgCl_2$, 0.2 M KCl pH 7.5) with ATP and L-methionine (10 mM final concentration of each), *Ec*MAT (0.4 mg mL^{-1} final concentration), *Ec*MTAN (0.025 mg mL^{-1}) and appropriate methyltransferase enzyme (10% v/v lysate for *Rn*COMT and 6OMT or 0.6 mg mL^{-1} final concentration of purified *Mx*SafC). Reactions were performed at 37 °C on a 200 μL scale. Reactions quenched by addition of 200 μL MeCN. Yields were determined by monitoring product formation against standards by analytical HPLC (method 1). Figure adapted from Roddan *et al.*[149]

For the *Rn*COMT reaction with the THIQ, **157**, a ratio of the two regioisomers C6-OMe:C7-OMe methylated products (**172:174**) were observed in an 87:13 ratio, as determined by 2D Nuclear Overhauser effect (NOESY) NMR spectroscopy. It was

therefore possible to conclude in the analytical RP-HPLC trace that the peaks at 7.2 and 7.3 min corresponded to the C6- and C7-*O*-methylated products, **172** and **174**, respectively. It was therefore assumed that both enzymes, 6OMT and *Mx*SafC were also regioselective towards the C6-position, and this was subsequently confirmed by NMR analysis of the purified product generated from the *Mx*SafC reaction. Low yields achieved with 6OMT meant that the products generated could not be isolated in sufficient quantities for structural determination. It is particularly interesting that C6-*O*-methylation is observed with MxSafC, as in previous reports, it is shown to be predominately regioselective towards the C7-hydroxyl, although high C6-hydroxyl selectivity has been observed for other substrates, as discussed further in Chapter 4.[209] The side-chain attached to the catechol clearly has subtle but effective influence in determining the position that methylation occurs. Therefore, using these bulky THIQs as substrates means that the substrate can only bind in a single orientation to the magnesium ion co-factor. Poor conversions were observed when using 6OMT lysate compared with using purified *Mx*SafC, likely due to a combination of poor enzyme expression and poor substrate acceptance. The lack of a hydrogen bonding interaction of the group at C-1 to key active site residue, Asp139 may also be contributing to low conversions due to decreased affinity for the substrate.

Preparative scale reactions were performed using the most productive methyltransferase, purified *Mx*SafC. As discussed in Chapter 4, attempts to perform the reaction with *Mx*SafC as clarified cell lysate were unsuccessful, likely because the enzyme is not particularly stable. Conversions (based upon calibration curves of purified product standards) when using 5 mM substrate, 20 mL reaction volume, were determined by RP-HPLC analysis of the two products (method 2) **171** and **172**, were found to be 33% and 46%, respectively. Enantiopurities of the products were assumed to be comparable to the starting material as the methylation step should not cause any racemisation.

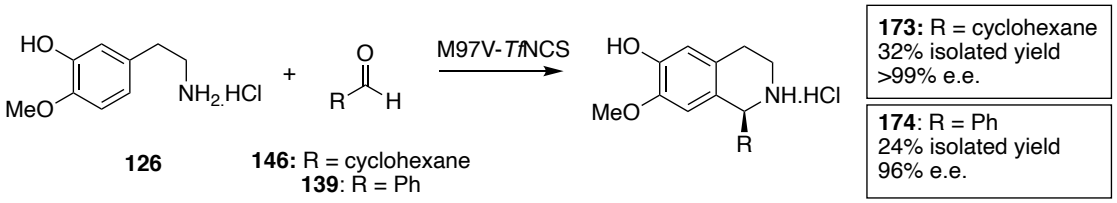
Interestingly, for reactions involving *Rn*COMT, altering the phenyl ring in the starting material (**157**) to a cyclohexane ring (**156**) instead resulted in exceptional regioselectivity, with only the *meta*-methylated product was observed (Scheme 5.4b). Attempts to rationalise these results using computational docking were unsuccessful and docking results contradicted what was observed experimentally.

5.2.8 Routes to single isomer, *para*-methylated (*S*)-1-aryl-THIQs

Accessing the analogous C7/*para*-methylated regioisomers of **171** and **172**, **173** and **174**, respectively, is valuable to generate analogues for drug discovery purposes. The *meta*-hydroxyl of dopamine is essential for a successful NCS reaction, whereas previous reports have shown that the *para*-hydroxyl is not, with catecholamines such as *meta*-tyramine and metaraminol being well-tolerated as substrates.[21,148] Although *Mx*SafC is capable of generating 4-methoxytyramine with high regioselectivity, conversions of 40% were achieved and isolating the pure product would be challenging, requiring preparative-HPLC.[209] Instead, 4-methoxytyramine (**126**) was generated chemically in a four-step synthesis, as given in Scheme 4.11 and discussed in Chapter 4.

NCS reactions were performed between **126** and aldehydes, **146** and **139** to give products (**173** and **174**, Scheme 5.5). With both aldehydes, RP-HPLC analysis showed complete consumption of the amine starting material. Isolated yields of (*S*)-**173** and (*S*)-**174**, after purification by preparative-HPLC (method 8), were 32% and 24%, respectively. The corresponding racemic products (*rac*-**173** and *rac*-**174**) were obtained by performing phosphate-catalysed reactions between **126** and **146** or **139** using standard conditions. The resulting products were also isolated by preparative HPLC and used for chiral HPLC analysis (method 5). High enantiomeric excesses were

achieved in both NCS catalysed reactions, >99% and 96% e.e., for (*S*)-**173** and (*S*)-**174**, respectively. Chiral HPLC analyses are given in the Appendix.



Scheme 5.5: Reactions towards C7-*O*-methylated (1*S*)-1-aryl-THIQs. 4-methoxytyramine (**126**) was used as the phenethylamine substrate in a M97V-*Tf*NCS mediated reaction.

5.2.9 Determining the selectivity of the catechol *O*-methyltransferases

The substrates used for the regioselective methylation reactions (**156** and **157)** were of high enantiopurity, however it is unknown whether the methyltransferases used are selective for one enantiomer of substrate. It would be advantageous if so, as this would improve the enantiopurity of the resultant products and would mean that if racemic substrate was used, the opposing enantiomer of starting material to the product generated could be isolated. Generating racemic standards of the C6-methylated products, *rac*-C6-OMe-**171** and *rac*-C6-OMe-**172** is labour intensive, involving the synthesis of 3-methoxytyramine followed by an acid-mediated Pictet-Spengler reaction and preparative HPLC purification.[242] Instead, reactions were performed with racemic substrates, *rac*-**156** and *rac*-**157** and activities compared to those using the analogous (*S*)-THIQs. Reactions involving *Mx*SafC as the methyltransferase did not go to completion and were selective towards the C6 position (Scheme 5.4). Therefore, to determine whether there was a preference for one enantiomer of substrate, once the methyltransferase reaction had been performed, the starting materials were reisolated by preparative HPLC and analysed by chiral HPLC. Racemic standards of C7-*O*-methylated THIQs, *rac*-**173** and *rac*-**174**, generated previously (Scheme 5.5) were used to aid determination of the *Rn*COMT reaction selectivity as some C7-*O*-methylated product is generated (Scheme 5.4). Analytical RP-HPLC analyses of the final reaction

mixtures are given in Figure 5.15 and chiral HPLC analyses of the reaction products are given in Figure 5.16.

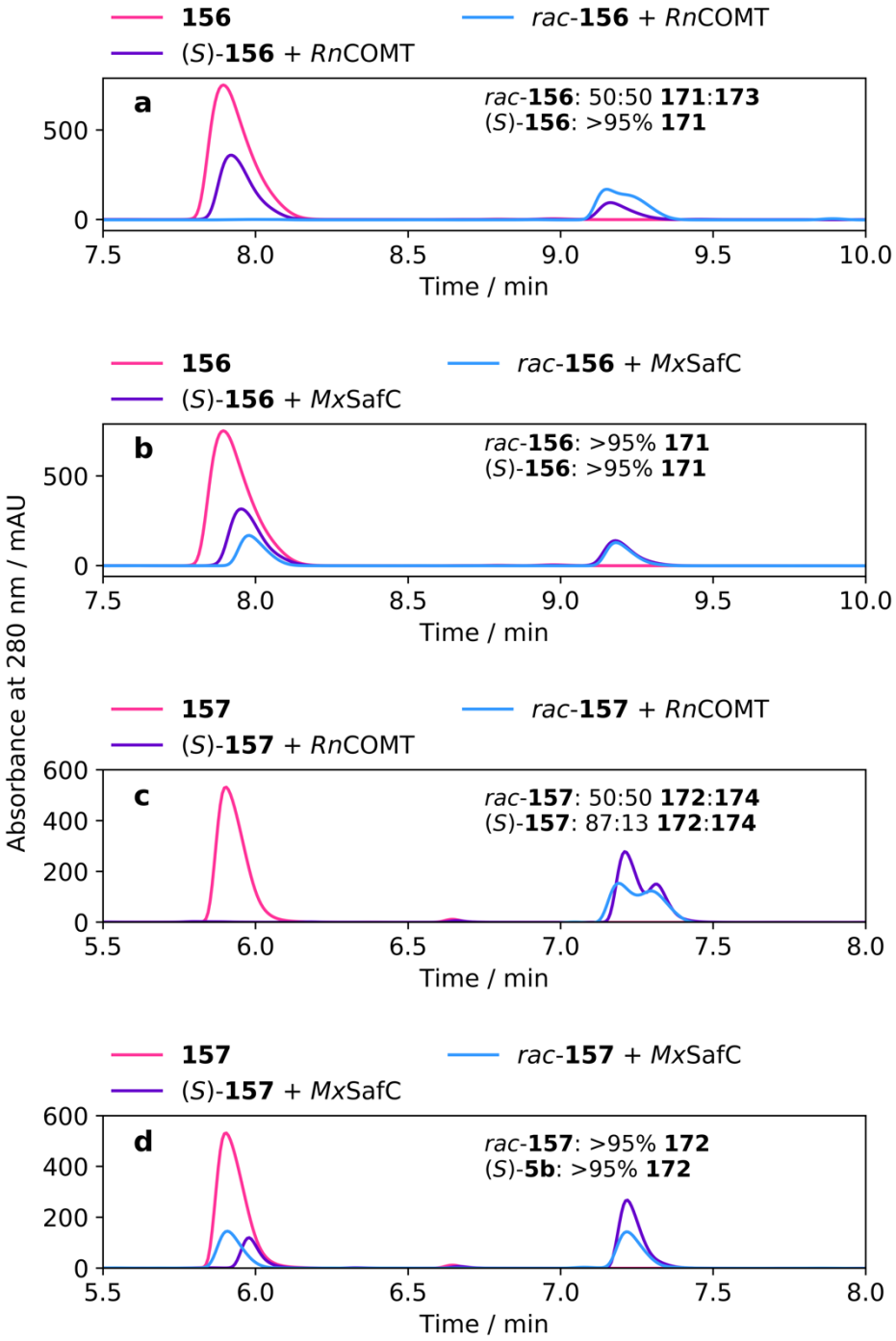


Figure 5.15: RP-HPLC analysis of methyltransferase-catalysed reactions. Reactions used **5a** or **5b** (either as the racemic or *S*-enantiomer, generated by a M97V-*Tf*NCS reaction) as the substrate reaction. In all cases, traces for **156** and **157** correspond to the purified, isolated starting material. Reactions were then performed between *rac*/*S*-**156** and **157** and *Mx*SafC or *Rn*COMT. RP-HPLC analysis of the final reaction mixture is shown, given with the ratios of C6:C7 *O*-methylated products generated. a) Reactions between **156** and *Rn*COMT. b)

Reactions between **156** and *Mx*SafC. c) Reactions between **157** and *Rn*COMT. d) Reactions between **157** and *MxSafC.* Figure adapted from Roddan *et al.*[149]

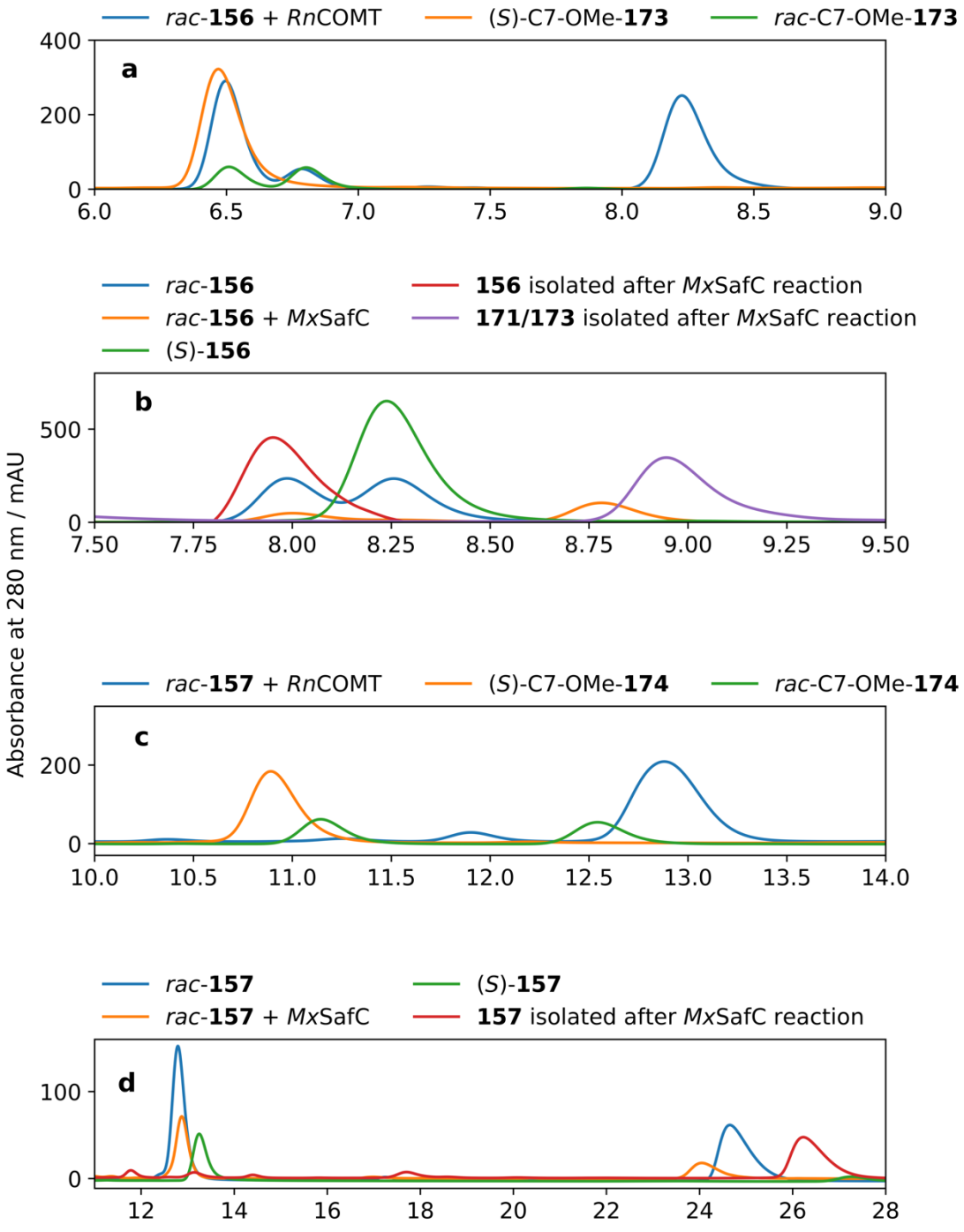


Figure 5.16 Chiral HPLC analysis of methyltransferase reactions. Reactions were performed between **156** and **157** (racemic) and methyltransferases, *Rn*COMT and *Mx*SafC. Resultant reaction mixtures or isolated products were analysed by chiral HPLC method 5. Products are compared to product standards. a) Isolated reaction product **171/173** obtained from *Rn*COMT catalysed methylation of *rac*-**156** compared with racemic and *S*-enantiomer standards of C7-OMe-**173**. b) Analysis of the methylation of **156** by *Mx*SafC; blue – standard of

rac-**156**, yellow – reaction mixture after methylation of *rac*-**156**, green, standard of *S*-**156**, red – isolation of **156** after *Mx*SafC methylation, purple – reisolation of **171** generated by *Mx*SafC-catalysed methylation of **156**. c) Isolated reaction product **172/174** obtained from *Rn*COMT catalysed methylation of *rac*-**157** compared with racemic and *S*-enantiomer standards of C7-OMe-**157**. d) Analysis of the methylation of **157** by *Mx*SafC; blue – standard of *rac*-**157**, yellow – reaction mixture after methylation of *rac*-**157**, green – standard of *S*-**157**, red – **157** isolated after *Mx*SafC reaction. Figure adapted from Roddan *et al.*[149]

For the methylation of **156** by *Rn*COMT (Figures 5.15a and Figures 5.16a), complete conversions were observed with both single enantiomer and racemic starting materials. With *rac*-**156**, a 50:50 product ratio of C6:C7 methylation was observed. Chiral HPLC analysis of the C7-methylated product **171** generated with this substrate, **156** compared with the racemic standard generated previously, *rac*-**171** (Scheme 5.4), showed that 84:16 (*S*):(*R*) product was generated, meaning that *Rn*COMT is neither regio- nor enantioselective towards *rac*-**156**. As described earlier, only C6-*O*-methylation is observed with (*S*)-**156** as the substrate, however the lack of selectivity observed with the racemic starting material means that using this methyltransferase will not improve enantiopurities in the products generated.

For the methylation of **156** by *Mx*SafC with either *rac*-**156** or (*S*)-**156** as the substrate, both reactions were regioselective giving only C6-OMe-**171.** Improved conversions were observed using racemic starting material (46% vs. 33% using (*S*)-**156**) (Figure 5.15b). To determine the enantioselectivity of the reaction, *rac*-**156** was used as the substrate and once the reaction was complete, the remaining starting material was isolated by preparative HPLC and analysed by chiral HPLC (Figure 5.16b). This demonstrated that only (*R*)-**156** remained, so *Mx*SafC is selective towards (*S*)-5a.

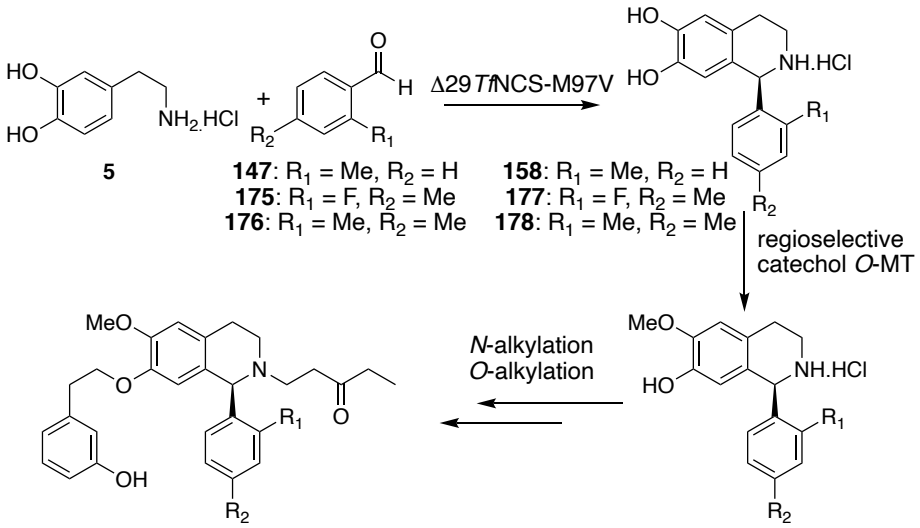
For the methylation of **157** by *Rn*COMT, complete conversions were also observed when using racemic starting material, *rac*-**156**. A 50:50 product ratio (**172:174**) for C6:C7 methylation was observed (Figure 5.15c) and no selectivity in the C7-methylation (to give C7-OMe-**174**) (Figure 5.16c).

For the methylation of **157** by *Mx*SafC, regioselectivity was observed with both racemic and single enantiomer starting materials, *rac*-**157** and (*S*)-**157** (Figure 5.15d) with similar conversions observed, 50% vs. 46% respectively. Chiral HPLC analysis of the reisolated starting material after reaction with *rac*-**157** indicated that again, the reaction was selective towards (*S*)-**157** (Figure 5.16d).

To conclude, with both racemic substrates (**156** and **157**) *Rn*COMT exhibited neither enantio- nor regioselectivity whereas *Mx*SafC was both enantio- and regioselective.

5.2.10 Reactions of disubstituted benzaldehyde derivatives with NCS

A variety of THIQs (of the general structure given in Scheme 5.6) are under patent due to high selectivity and inhibitory action towards RAS proteins (HRas, KRas and NRas) and in particular, the KRas proteins.[244] The genes encoding for these proteins are the most common oncogenes in human cancers (nearly 30% of all human tumours). Mutations in these proteins causes permanent activation resulting in overactive signalling and increased cell growth and division, thus giving rise to cancers.[244] Finding inhibitors of these proteins is challenging as there are no deep, hydrophobic pockets in which to bind small-molecules.[245] Therefore, identifying selective inhibitors of these proteins is important in the ongoing search for novel, effective small-molecules for cancer treatments.

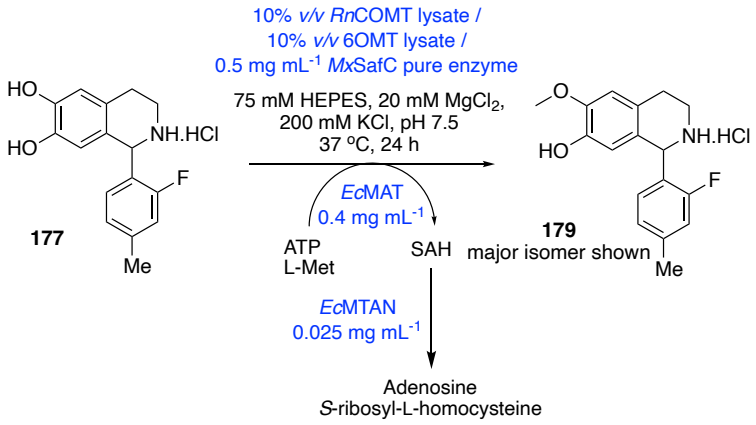


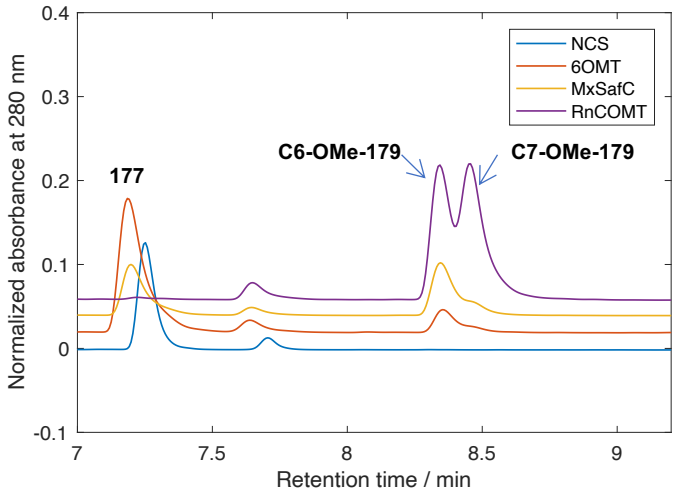
Scheme 5.6: KRas inhibitor THIQs that could feasibly be generated using an NCS-mediated synthesis followed by a regioselective *O*-methyltransferase reaction.

In the patent, the THIQs are synthesised using a protecting-group based strategy, with the ring closure performed via an acid-mediated Pictet-Spengler reaction or a Bischler-Napieralski reaction, followed by a reduction to give the THIQ scaffold. The stereochemistry at C-1 is not defined and only procedures to give the racemic compounds are described. It is assumed that the stereocentre would be defined during an asymmetric hydrogenation step as detailed previously.[244]

Therefore, to develop the THIQ scaffold of these small molecules and selectively synthesise the (*S*)-enantiomer of product, it was hoped that an NCS reaction could be performed between dopamine and a relevant disubstituted benzaldehyde analogue (**175** or **176**) to give THIQs, **177** and **178**. The selective methylation of one of the catechol C6 hydroxyls could then be performed using one of the regioselective catechol *O*-methyltransferases. The addition of other groups on the nitrogen and oxygen could then be performed sequentially afterwards. This would give a regioselective and stereoselective route to these compounds in four-steps compared with the 8 -11 step synthetic route described to give the racemic compounds.

NCS-mediated reactions were performed between dopamine and two disubstituted benzaldehydes; **175** and **176**. Reactions were not performed with 2-chloro,4-methylbenzaldehyde, which would be a route to another desirable THIQ product, due to poor enantioselectivities in the products with a single halogen. Low conversions were observed with **175** (Scheme 5.7), as the aldehyde substrate, with a minimal background reaction observed in the control sample. No conversions were observed for reactions with **176**, likely due to the bulky nature of the substrate. Representative HPLC of traces of both reactions with background traces are given in Scheme 5.7.





Scheme 5.7: Achiral HPLC traces of methyltransferase reaction of 177 to give 179 using a variety of regioselective *O*-methyltransferases. *Reaction conditions:* Once the NCS reaction was completed (performed on 100 μL scale), the reaction mixture was quenched and lyophilised (workup method 5). The residue prepared in 25 mM HEPES, 20 mM $MgCl_2$, 0.2 M KCl pH 7.5) with ATP and L-methionine (10 mM final concentration of each), *Ec*MAT (0.4 mg mL^{-1} final concentration), *Ec*MTAN (0.025 mg mL^{-1}) and appropriate methyltransferase enzyme (10% v/v lysate for *Rn*COMT and 6OMT or 0.6 mg mL^{-1} final concentration of purified *Mx*SafC). Reactions were performed at 37 ºC on a 200 μL scale. Reactions quenched by addition of 200

μL MeCN and analysed by HPLC method 2. Yields were determined by monitoring product formation against standards by analytical HPLC (method 1).

The enantiopurity of the resultant enzymatic product (**179**) was determined by chiral HPLC (method 5) but unfortunately, the observed product was racemic, suggesting that a racemic reaction is associated with the enzyme sample. Regioselective methylation reactions were attempted with *Rn*COMT, 6OMT and *Mx*SafC using racemic **177** as the THIQ substrate, as shown in Scheme 5.7.

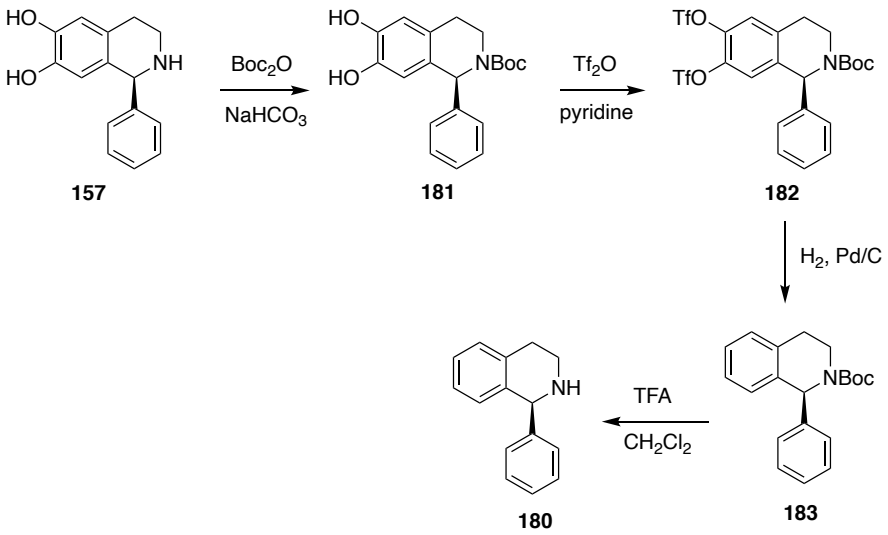
Complete consumption of the starting material was observed with reactions with *Rn*COMT, however no regioselectivity was observed. It appeared that a 50:50 mixture of the C6:C7 *O*-methylated product (**179**) was formed. 6OMT and *Mx*SafC were most regioselective, generating product in 4:1 and 3:1 ratio of C6:C7 *O*-methylated products, respectively. This assumes that as with products **171** and **172**, the C6-OMe product elutes first on aHPLC. It is possible that *Mx*SafC is selective towards the (*S*)-enantiomer of **177** but further analysis as in Figures 5.12 and 5.13 would need to be performed to determine this. Regardless, the regioselective methylation of the racemic THIQ is still synthetically useful.

5.2.11 Removal of catecholic hydroxyl groups

There would be two advantages of removing the catecholic hydroxyl groups from the THIQ products generated by NCS-mediated reactions. Firstly, the catechol is highly oxidatively sensitive and can undergo a radical polymerisation in the presence of oxygen. Removing the hydroxyl groups would therefore improve the stability of the final compounds generated and expand the portfolio of (*S*)-THIQs that could be generated using an NCS-based synthesis. One of the major advantages of using a Pictet-Spenglerase is the ability to generate products in high enantiomeric excess and high yield in a single step. This, combined with the wide substrate scope of NCS, would result in the generation of a wide variety of THIQs in high enantiopurity.

In particular, this would lead to the synthesis of the THIQ precursor (**180**) to solifenacin succinate (**81**), a widely used anti-muscarinic for the treatment of an overactive bladder.[246] Industrial syntheses involve the synthesis of the racemic compound via Bischler-Napieralski methods followed by resolution of the two enantiomers using L-tartaric acid to give the product in high enantiomeric excess (>99%). The (*R*)-enantiomer of product remaining in the mother liquor can be then racemised by refluxing in KOH/DMSO and the recrystallisation procedure repeated.[247] A major advantage of an NCS-mediated route would be the ability to generate analogues of solifenacin, varying substituents on the C-1' aromatic rings.

The stereochemistry of all products generated by an NCS-mediated reaction has thus far been assumed to always give the (*S*)-centre at C-1. However, it is possible that the stereochemistry of the products can switch with different substrates. Another Pictet-Spenglerase enzyme, strictosidine synthase (STR), has substrate dependent selectivity. When using small, aliphatic aldehydes, to give (*R*)-products are generated, rather than when the bulky natural aldehyde substrate, secologanin, is used, (*S*) stereochemistry at C-1 is generated.[89] This is due to the unusual inverted binding of the substrates in the enzyme active site when reactions are performed with smaller, aliphatic aldehydes.[90] Therefore, with NCS, there is a possibility that using benzaldehydes as substrates will change the stereochemistry of the products and so it is important to absolutely determine the chirality at C-1 rather than assume that it is (*S*) by literature precedent. As solifenacin succinate is routinely synthesised, chiral HPLC data and x-ray crystallography data are known for (1*S*)-1-phenyl-1,2,3,4-THIQ (**180**). It was hoped that this molecule could be synthesized via the removal of the catechol hydroxyl groups of the THIQ product formed from the NCS reaction between dopamine (**5**) and benzaldehyde as demonstrated in Scheme 5.8.



Scheme 5.8: Proposed synthetic route for removal of catecholic hydroxyl groups of THIQ products of the NCS reaction between dopamine and benzaldehyde.

The synthesis was performed without isolation of each intermediate due to the expected instability of each. The presence of each intermediate was identified by low resolution mass spectrometry (LC-MS). There are no reported syntheses to remove catechol hydroxyl groups to give benzene, nor any other analogous syntheses. The removal of a single phenolic hydroxyl is however, widely reported in the literature, *via* the activation of the hydroxyl group to a triflate or mesylate, followed by hydrogenation to cleave the carbon-oxygen bond neighbouring the aromatic ring.[248] It was therefore hoped that these synthetic methods could be used to remove the phenolic hydroxyls.

The THIQ starting material, **157** was obtained by an NCS-mediated reaction between dopamine (**5**) and benzaldehyde (**139**) (0.2 mg mL^{-1} final concentration Δ29*Tf*NCS-M97V) and semi-purified via an extractive workup (workup method 3) to remove any unreacted dopamine. Some **139** remained but would not interfere with subsequent reactions. As triflic anhydride can be attacked by both oxygen and nitrogen nucleophiles, the THIQ nitrogen in **157** was Boc protected. The resulting product (**181**) was isolated via extractive workup. Some Boc anhydride remained however it was deemed not to interfere with subsequent synthetic steps and could be easily removed later in the synthesis when the Boc-amine was cleaved. Purification of **181** would also

be challenging due to chelation of the catechol to silica. The crude product **181** was therefore taken through to the next step of synthesis without further purification.

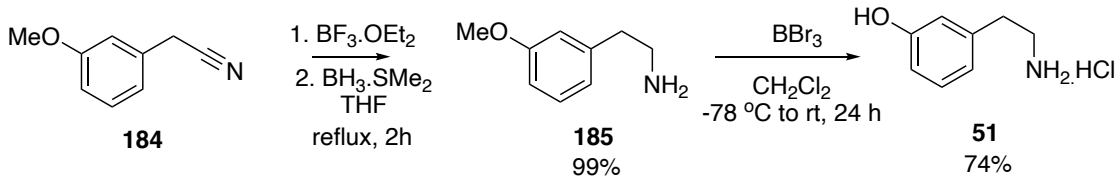
The method used for triflate protection is widely reported. The reaction involves the deprotonation of the hydroxyl groups with base, most commonly triethylamine or pyridine. This improves the nucleophilicity of the hydroxyl group and is followed by addition to triflic anhydride to give the triflate-protected product. It was confirmed by LC-MS that triflate protection of **181** to give **182** had occurred, however the Boc-protecting group was not observed. It was unknown whether the Boc group is being removed during the mass spectrometry ionisation process or whether it is being cleaved during the addition of triflate as the conditions are highly acidic. Both di- and tri-triflate substitution was observed. This was deemed not to interfere with subsequent steps as triflate is commonly used as a nitrogen protecting group and during hydrogenation, the group would be removed to regenerate the amine. Hydrogenation to give **183** was attempted using 10% palladium on carbon as the catalyst and tetrahydrofuran as the reaction solvent. No desired product was observed, only starting material. The reaction was also performed with adding a small amount of methanol or formic acid to the solvent, and by bubbling hydrogen through the solvent but again, only starting material was isolated. It was assumed that the two triflate groups were stabilising each other and it was hoped that another catalyst may be suitable. A paper by Miller *et al.*[249] used 10 mol % $Pd(dppf)Cl_2$ catalyst with formic acid and triethylamine to remove a single triflate from a conjugated naphthalene system with methoxy and acetyl groups presents, however again, no desired product was observed by LCMS.

It was thought that perhaps the difficultly in triflate removal was due to the strongly electron withdrawing aspects of the triflate groups and that the neighbouring triflate groups were in some way stabilising each other. Therefore, the synthesis was repeated with mesylate groups rather than triflate, however no product was observed when using

both methods of hydrogenation (Pd/C and Pd(dppf)Cl_2). More harsh hydrogenations conditions may be required for the carbon-oxygen bond cleavage and high-pressure hydrogenation may be successful in giving the desired product. To avoid the use of high-pressure reaction conditions, the same procedure was attempted but *meta*-tyramine (**51**) was used as the amine substrate in the NCS reaction rather than dopamine. This would result in a THIQ with a single phenolic hydroxyl group, and it was hoped that this could yield the product as the hydrogenation reaction should proceed following literature precedent.

5.2.12 Removal of a single THIQ phenolic hydroxyl group

Meta-tyramine (**51**) is commercially available, however it is expensive (*ca.* £200 for 1 g) and since there is a widely-reported synthetic route (Scheme 5.9), it was synthesised in a two-step synthesis starting from 2-(3-methoxyphenyl)acetonitrile (**184**).[22] The first step of the synthesis involves the reduction of nitrile (**184)** to the amine (**185**), which can be purified *via* an extractive acid-base work-up. The final step of the reaction is the demethylation to give *meta*-tyramine (**51)**. The reaction went to completion (as monitored by TLC); however, some material was lost in the workup procedure. The final product (**51**) was purified by recrystallisation in *n*-hexane.



Scheme 5.9: **The synthetic route performed to give *meta*-tyramine (51) in 74% yield**.[22]

Using *meta*-tyramine (**51**) as the amine substrate for NCS reaction often results in significantly lower conversions (*ca.* 50% of those observed with dopamine instead) (Table 3.1). Indeed, in ligand-bound structures of NCS with non-productive intermediate mimics with the dopamine catechol, both hydroxyl groups of the catechol hydrogen-bond to Lys-122. It is known that the *para*-hydroxyl group is not essential for a

successful NCS reaction, but the *meta*-hydroxyl is. However, the additional *para*-hydroxyl group may make the reactions more productive by providing extra 'anchoring' in the active site or assist with the *meta*-phenol deprotonation.

To find optimal reaction conditions for gaining a high enantiomeric excess in the product **186** of NCS reactions between *meta*-tyramine (**51**) and benzaldehyde (**139**), a timepoint assay was performed (Figure 5.17), analogous to those performed for the reactions between dopamine and benzaldehyde or 3-methylbenzaldehyde (Figures 5.7 and 5.8). For all reactions with dopamine as the amine substrate, the highest e.e.s were observed with a high concentration of enzyme (0.5 mg mL^{-1} final concentration). Therefore, here all reactions were performed with this concentration of enzyme but instead, the different NCS mutants that were previously shown to be promising with α-substituted aldehydes (namely M97F and M97V) were used as the reaction catalysts (Figure 5.16), in case altered activities were observed when changing the amine donor. For reactions performed for 24 h using Δ29*Tf*NCS-M97V as the reaction catalysts gave complete conversion to the product, which was formed in 94% e.e. Therefore, these conditions were used on a preparative scale to give the THIQ product (**16**).

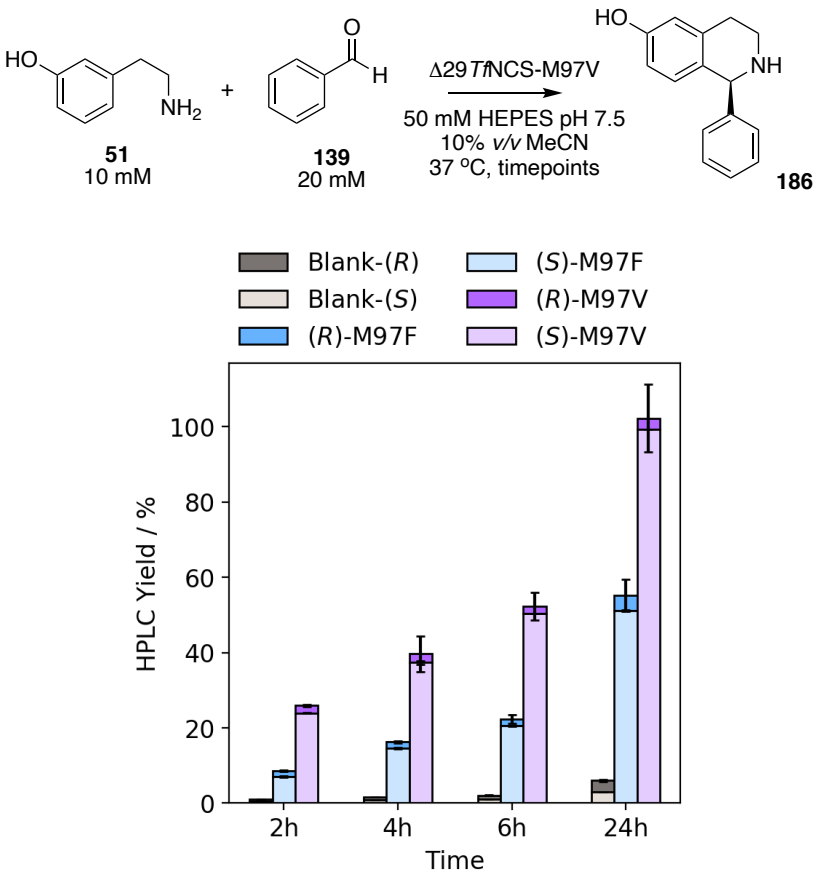
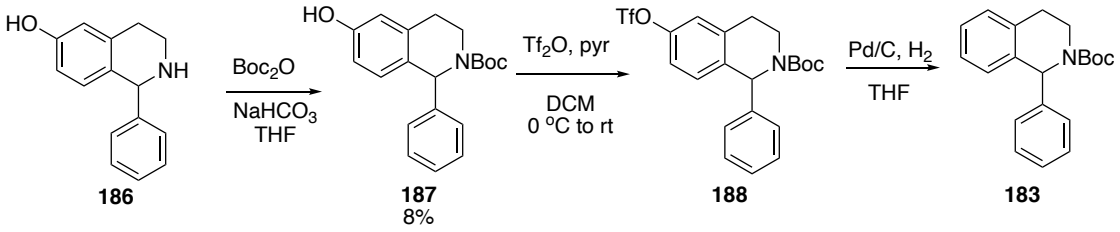


Figure 5.17: Timepoint assay of the NCS-mediated reaction between *meta*-tyramine (51) and benzaldehyde (139). Conversions and enantiomeric excesses in the resulting products (**186**) were messaged. *Reaction conditions*: 10 mM dopamine HCl, 10 mM sodium ascorbate, 20 mM benzaldehyde in HEPES buffer (50 mM, pH 7.5) with 10% *v/v* MeCN and final concentration of *Tf*NCS at 0.2 or 0.5 mg mL^{-1}. Control reactions were performed using the same conditions but the *Tf*NCS sample was substituted for an equal volume of enzyme buffer (20 mM Tris, 50 mM NaCl, pH 7.5). Reactions were performed at 37 °C for 3 h, 100 µL scale reactions; samples were prepared by workup method 1, yields were determined by monitoring product formation against standards by analytical achiral HPLC (method 1).

The synthetic route performed for the removal of the phenolic hydroxyl group was then performed in a similar manner (Scheme 5.10), with the same rationale to that attempted for the catechol hydroxyl groups (as shown in Scheme 5.9). Due to potential instability of the intermediates, only the Boc-protected amine (**187**) was isolated, so yields were not determined at each subsequent step.

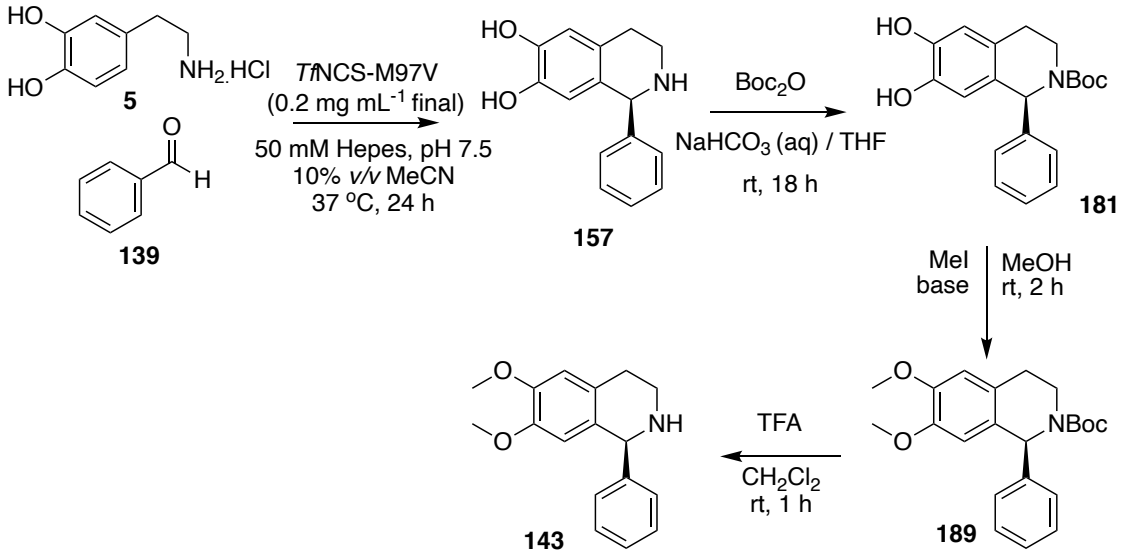


Scheme 5.10: **Synthetic route attempted for the removal of a single THIQ phenolic hydroxyl group.**

The product of the NCS reaction (**186**) was purified via extractive workup to remove any remaining meta-tyramine or benzaldehyde. To improve the Boc protection reaction, the hydrochloride salt was resuspended in saturated sodium hydrogen carbonate and extracted into ethyl acetate to give the free amine (**186**). The Boc protection was performed as usual, and monitored by TLC, although it was challenging to determine whether the reaction had gone to completion. The product was isolated via extractive workup then further purified by column chromatography. A very low yield of product **187** was isolated (8%). The conversion of the triflate protection of **187** to give **188**, was also poor. Therefore, as this synthesis proved challenging and would not be a feasible route to give stereoselective THIQ products without hydroxyl groups. Instead, to determine the stereochemistry at the C-1 position, another synthetic route was attempted based on compounds found in the literature with determined stereochemistry and reported chiral HPLC conditions.[250]

5.2.13 Determination of the stereochemistry of (1*S*)-aryl THIQ products generated by *Tf*NCS

Chiral HPLC and optical rotation data is available for the methoxy protected analogue of **157**, **143**.[250] As nitrogen is more nucleophilic than the catecholic hydroxyl groups, the product was synthesised via the Boc protection of the amine (as shown in Scheme 5.11).



Scheme 5.11: **Synthetic route to give (*S*)-6,7-dimethoxy-1-phenyl-1,2,3,4-tetrahydroisoquinoline (20) via a stereoselective NCS-mediated Pictet-Spengler reaction.**

Based upon timepoint optimisation reactions, the NCS reaction was performed for 24 h with 0.2 mg mL^{-1} final concentration of Δ33*Tf*NCS-M97V (Figure 5.14). The resultant product was purified by extractive workup (method 3) to remove any remaining dopamine or benzaldehyde. The workup method was adapted in an attempt to improve isolation yields. The workup procedure was monitored by analytical HPLC (method 1) to ensure that a minimal amount of product was lost. After the NCS reaction was performed, the reaction mixture was directly extracted into ethyl acetate, without quenching of the reaction. The product was dried and resuspended in 1 M HCl. This aqueous sample was analysed by RP-HPLC and any remaining dopamine had been removed but still some benzaldehyde remained which was removed via extraction with dimethyl carbonate. The aqueous layer was monitored by RP-HPLC to check for benzaldehyde removal. It was found that a single dimethyl carbonate wash was sufficient to remove the excess aldehyde, rather than four as previously reported.[118]

The product (**157**) was isolated, of high purity and in reasonable yield (60%) and was Boc-protected using a standard literature procedure.[251] The reaction was performed overnight with an excess of Boc anhydride to try to mitigate previous issues of poor

conversions. The consumption of the starting material was monitored by TLC analysis and once completed; **157** was isolated via extractive workup. Due to the presence of the catechol and issues with chelation to silica, purification of intermediate, **181** is challenging, therefore the product was taken through crude to the subsequent methylation step. The methylation reaction is analogous to that used in Chapter 4, where the addition of potassium carbonate was sufficient to deprotonate and therefore methylate the catechol hydroxyls. The product (**189**) was isolated by silica column chromatography to yield 6.5 mg. The yield was low likely due to the recurring issue of poor Boc protection. The product was characterised by HRMS, with *m/z* values of 370.30, 314.14 and 270.50 observed, corresponding to masses of the product (**189**), loss of the *tert*-butyl group and Boc group, respectively.

The Boc deprotection of the product to give **143** was performed using a standard procedure. The reaction completed in 5 minutes as monitored by TLC analysis. The resulting product was purified by semi-preparative HPLC as the product was present in very low quantities. Chiral HPLC was performed using the same conditions reported by Evanno *et al*.[250] (chiral HPLC method 6). The two enantiomers have a retention times: (*S*) = 17.8 min, (*R*)= 25.7 min. As these conditions were already reported, racemic **143** was not synthesised. Two major peaks were observed, at 16.2 and 24.3 min i.e., both eluting 1.4 min earlier than those reported. This is not uncommon, with slight differences in column age, temperature and mobile phase composition having significant effects on retention time with chiral, normal phase columns. However, minor impurities were observed in chiral HPLC analysis which overlaps with peaks corresponding to the desired product. To further separate the peaks, the polarity of the mobile phase was decreased using 5% isopropanol, rather than 10% isopropanol (method 7). Chiral HPLC analysis of **143** is given in Figure 5.18, with Figure 5.18a corresponding to conditions given in method 6 and Figure 5.18b corresponding to conditions given by method 7. The polarity of the mobile phase was decreased to improve separation of peaks.

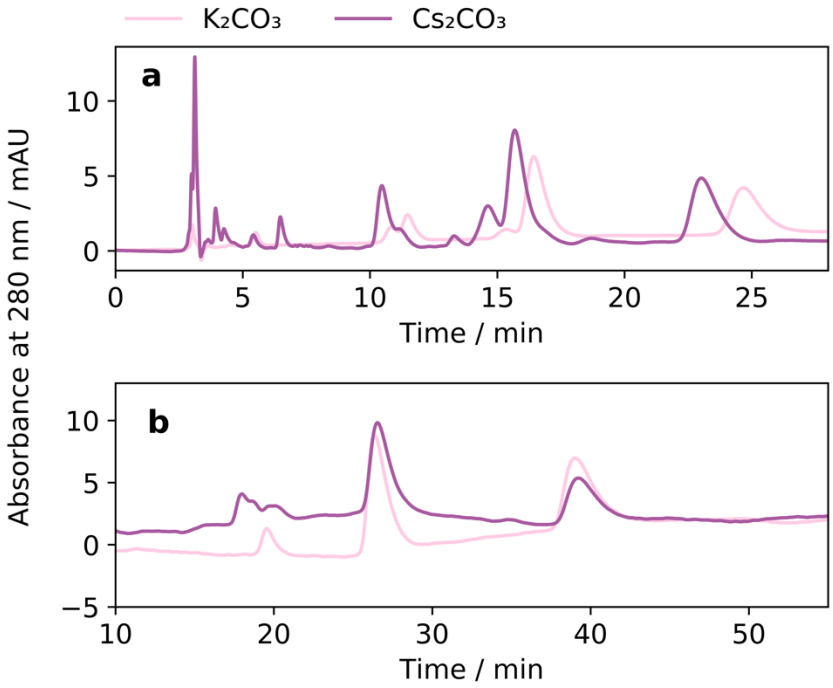


Figure 5.18: Chiral HPLC analyses of 143. a. Using a mobile phase of 10:90:0.1 isopropanol:*n*-hexane:diethylamine t_R (*S* enantiomer) = 16.2 min, 564.2 mAU^2, t_R (*R* enantiomer) = 24.3 min, 499.6 mAU^2. **b**. Using a mobile phase of 5:95:0.1 isopropanol:*n*-hexane:diethylamine t_R (*S* enantiomer) = 26.5 min, , t_R (*R* enantiomer) = 39.2 min.

Using potassium carbonate as the base in the methylation step (**181** to **189**) gave an enantiomeric excess of 6% in the product, with (*S*)-**146** as the major enantiomer. Therefore, significant racemisation must have occurred during the chemical modification of **181**. The C-1 proton is highly acidic, due to the neighbouring electron withdrawing Boc-amine and the phenyl groups, so is likely that the basic conditions used to deprotonate the catechol hydroxyl groups, potassium carbonate in methanol, were capable of deprotonating at the C-1 position and thus causing racemisation.

To improve the enantiomeric excesses in the products, the methylation step was attempted using two bulkier bases, caesium carbonate or 2,6-lutidine. It was hoped that this would reduce deprotonation at the sterically hindered C-1 position. A large excess of iodomethane (20 e.q.) was also used in an attempt to reduce the reaction time. No conversion to give **189** was observed using 2,6-lutidine as a base, likely due to the

inaccessibility of the nitrogen lone pair. The reaction involving caesium carbonate proceeded successfully to give **189**, and the product was taken through to the next step (Boc deprotection) crude and isolated by preparative-HPLC (method 4). Chiral HPLC analysis (by methods 6 and 7) gave a 20% e.e. of the final product (**143**). Racemisation is clearly occurring due to the presence of base but as an excess of the (*S*)-enantiomer is observed, the stereochemistry of the *Tf*NCS reaction with aldehydes is confirmed.

5.2.14 Structural evidence for the acceptance of the benzaldehydes by norcoclaurine synthases

Despite reports that benzaldehydes were not accepted as substrates by *Tf*NCS, the first attempted co-crystallised structure of *Tf*NCS was gained by Ilari *et al*. in 2009, with the ligands, dopamine (**5**) and 3-hydroxybenzaldehyde (**168**).[27] Analysis of the co-crystallised structure shows weak positive electron density in the active site, suggesting low occupancy and multiple orientations of the substrates[28,30]. Work presented in this chapter has demonstrated that 3-hydroxybenzaldehyde is in fact accepted as a substrate, albeit poorly. Therefore, as crystals can retain enzymatic activity, it is also possible that unclear density was observed because they were in fact observing a mixture of reaction transition states. Also, as low conversions are observed with this aldehyde as a substrate (Figure 5.6), it is likely that poor occupancy was also a contributing factor.

Into this positive electron density, the aldehyde was modelled to be close to the Lys-122 and the dopamine was modelled to be nearer the active site entrance, leading to an aldehyde-first mechanism being proposed. This was later disputed due to other chemical and steric reasons (as discussed in Lichman *et al.*[28] and in Chapter 1). It is also possible that the two substrates are capable of binding 'aldehyde-first' but a productive reaction only occurs when they bind 'dopamine-first'. All computational docking studies and the co-crystallised structures gained thus far have not been with

benzaldehydes and there is a possibility that different substrates interact with residues in alternative ways. For example, with strictosidine synthase (STR), changing the carbonyl substrate from the bulky glycosylated monoterpene, selcoganin (Section 1.7.1) to a variety of low molecular weight, aliphatic aldehydes changes the stereochemistry of the resultant tetrahydro-β-carbolines from (*S*) to (*R*) at the C-1 position.[89] A co-crystallised structure of *Op*STR with a non-productive reaction intermediate analogue bound in the active site, was gained by Eger *et al*., and inverted binding of the substrates was observed. It is therefore possible that a similar phenomenon may be observed with NCS when benzaldehydes are used as substrates, so gaining structural evidence for benzaldehyde acceptance was important.

To understand the acceptance of benzaldehyde and derivatives by *Tf*NCS, it was hoped that a co-crystallised structure could be gained with a non-productive intermediate mimic of the reaction iminium ion intermediate in the active site. Since improved conversions were observed with the M97V active site mutant, a crystallisable construct was developed i.e., using a His-tagged, truncated form of NCS, Δ33*Tf*NCS-M97V with a TEV-cleavage site for removal of the His-tag. Cloning of this construct was performed by John Ward and the construct was codon optimised for expression in *E. coli*. This construct had proved promising in previous crystallographic trials to give high resolution structures.[30,148] Expression and purification was performed as previously reported for the wild-type construct and pure protein was obtained in high yield (55 mg from a 1 L culture). The pure enzyme was concentrated to 12.6 mg mL^{-1} and stored at -80 °C in TRIS buffer (20 mM with 50 mM NaCl, pH 7.5), consistent with previous crystallisation trials performed with the wild-type enzyme. SDS-PAGE analysis of the purification process is given in Figure 5.19.

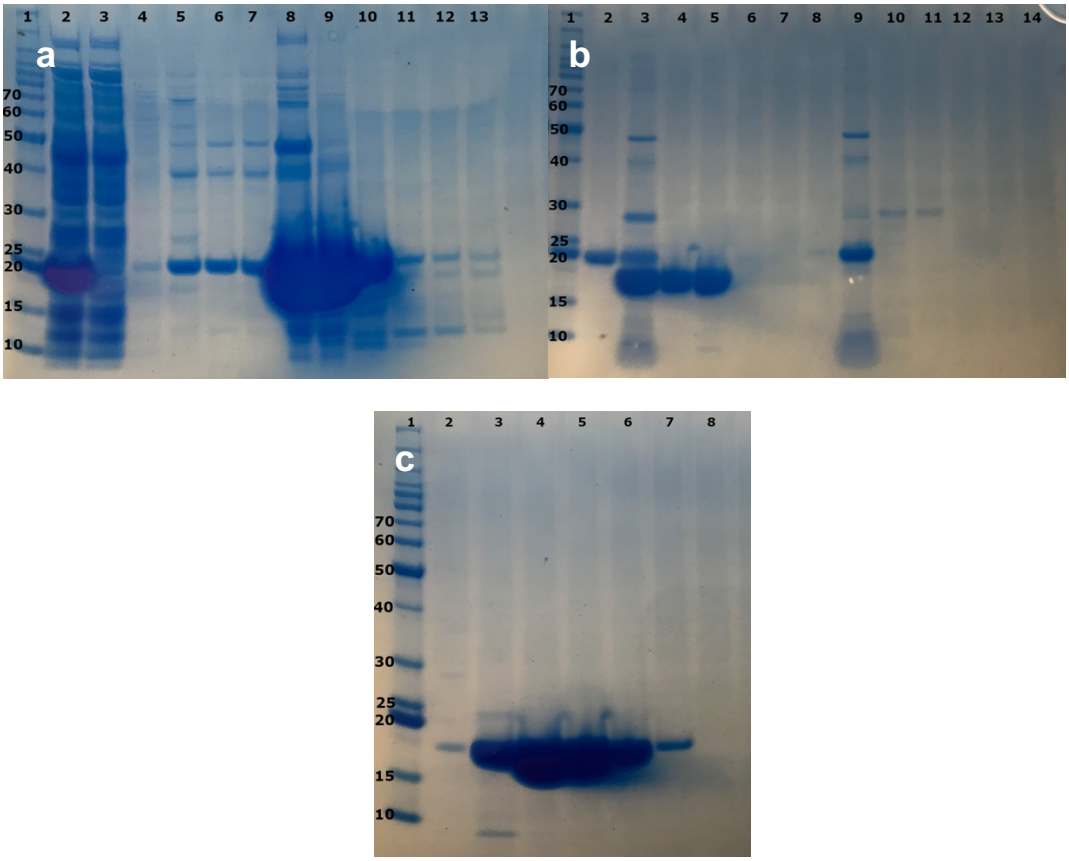
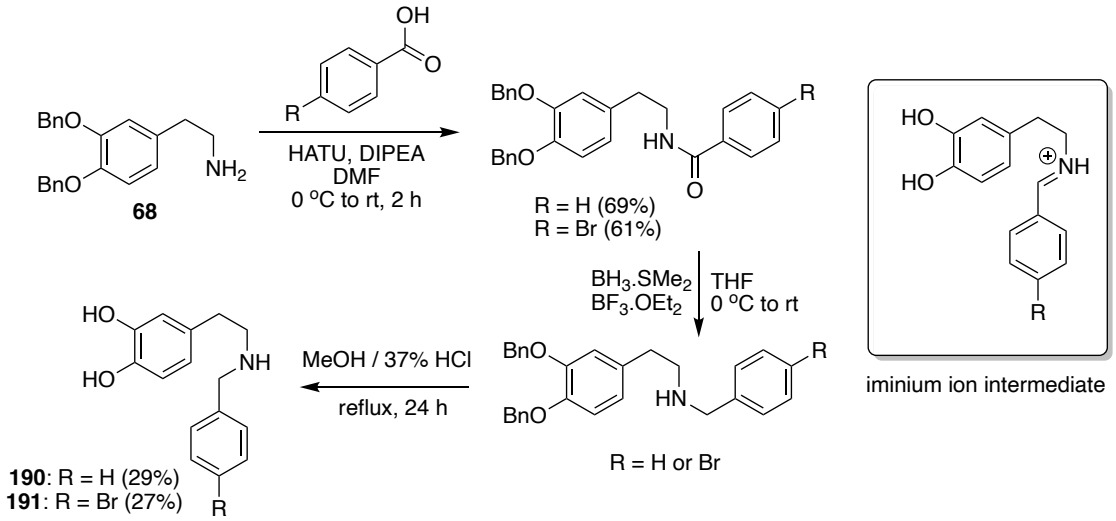


Figure 5.19: SDS-PAGE analysis of Δ33*Tf*NCS-M97V purification. a. His-trap purification of Δ33*Tf*NCS-M97V (protein expected at 20.7 kDa): Lanes: 1, Benchmark™ Protein Ladder (masses given in kDa) 2, clarified cell lysate loaded onto column. 3, flow through with lysis buffer. 4, wash with 6 CV 20 mM imidazole buffer. 5, wash with 6 CV 40 mM imidazole buffer. 6-13, wash with 500 mM imidazole buffer and collected 3 mL fractions. **b.** His-trap purification of Δ33*Tf*NCS-M97V after TEV cleavage to remove TEV protease. Lanes: 1, Benchmark™ Protein Ladder, masses given in kDa. 2, Pre-TEV cleavage sample (Lane 7 from gel 1) 3, sample loaded onto column. 4, flow through with lysis buffer. 5, 3 CV 20 mM imidazole. 6, 8 CV 20 mM imidazole. 7, 2 CV 40 mM imidazole. 8, 5 CV 40 mM imidazole. 9 and 10, 2 CV 500 mM imidazole. 11, 5 CV 500 mM imidazole. **c.** Gel filtration of Δ33*Tf*NCS-M97V: gel of peaks isolated from major peak from Superdex S75 gel filtration. Lane 1: Benchmark™ Protein Ladder, masses given in kDa. Protein samples corresponding to lanes 4-7 were pooled and used for crystallographic studies. Protein expected at 18.1 kDa.

Two different non-productive reaction intermediate mimics were synthesised, based upon the iminium ion intermediate of the reactions between **5** and benzaldehyde (**139**) or 4-bromobenzaldehyde (**170**). Both were synthesised using the same synthetic route the two previous co-crystallised structures of *Tf*NCS.[30,148] These two aldehydes were chosen because firstly, the greatest conversions with an aromatic aldehyde were

observed with benzaldehyde. Cyclohexanecarboxaldehyde (**146**) was not used due to a lack of electron density and high flexibility in the ring which could lead to challenges in determining ligand placement if multiple conformations and poor occupancy was observed (as was the case with attempted to co-crystallise other ligands as described in Chapters 2 and 3). In an attempt to overcome these limitations and increase the certainty of ligand presence in the enzyme active site, a mimic was synthesised based on the reaction between **5** and **170** as the identification of a bromine atom in the active site can typically be determined using anomalous diffraction methods.[252] Although **170** was poorly accepted as a substrate by *Tf*NCS (Figure 5.5), it was hoped that co-crystallisation might be observed. Both mimics were synthesised using an analogous synthetic route to other mimics synthesised (Sections 2.2.4, 3.2.6 and 6.2.4), as shown in Scheme 5.12.



Scheme 5.12: Synthetic route to give two benzaldehyde reaction intermediate mimics. The mimics are based upon the iminium ion intermediates of the Pictet-Spengler reactions between dopamine (**5**) and benzaldehyde (**139**) or 4-bromobenzaldehyde (**170**).

The synthesis of the mimics proceeded simply. Once isolated, both final products (**190** and **191**) were resuspended in DMSO to give a final concentration of 200 mM. Both were stored at -80 °C under argon as DMSO is a known oxidant and the molecules are oxidatively-sensitive as there is a free catechol.

Initial crystallisation screens were performed by the sitting-drop method using commercially available, 96-well screens; Structure, JCSG, PACT and Index. These screens have previously contained conditions which have resulted in *Tf*NCS crystallisation.[31] The protein was prepared for the following screens: apo protein (at 12.6 mg mL^{-1}), with benzaldehyde reaction mimic (**190**) (10 mM final concentration, 12 mg mL^{-1} protein) and 4-bromobenzaldehyde reaction mimic (**191**) (10 mM final concentration, 12 mg mL^{-1} protein). In the case of the benzaldehyde mimic, **190,** addition to the protein resulted in the temporary precipitation of the enzyme which was eventually resuspended in the buffer after a couple of minutes of mixing. Unfortunately, with addition of the 4-bromobenzaldehyde mimic, **191**, irreversible precipitation of the enzyme occurred, and crystallisation trials could not be performed.

For both apo and ligand-bound crystallisation trials, crystals were observed using many different conditions after one week. A range of crystals were sent to a synchrotron (Diamond Light Source, i04) for screening. The conditions of the crystals observed, and the resolution of the datasets collected are given in Table 5.1. The cryo-protectant used was 80% mother liquor, 20% ethylene glycol.

Table 5.1: Conditions in which protein crystals were observed from initial co-crystallisation screens with benzaldehyde reaction mimic, 190.

Dataset	*Protein*	*Salt*	*Buffer*	*pH*	*Precipitant*	*Dataset resolution / Å*
1	Apo	-	-	-	10% w/v PEG 1000 10% w/v PEG 8000	3.5
2	Apo	0.2 M NaCl	0.1 M Bis-Tris	5.5	25% w/v PEG 3350	3.4
3	Apo	0.1 M NH_4OAc	0.1 M Bis-Tris	5.5	17% w/v PEG 10000	3.6
4	Apo	50 mM KH_2PO_4	-	-	20% w/v PEG 8000	3
5	Holo	50 mM KH_2PO_4	-	-	20% w/v PEG 8000	2.2

6	Holo	1.0 M $NaOAc.3H_2O$	0.1 M Imidazole	6.5	-	2.8
7	Holo				10% w/v PEG 1000 10% w/v PEG 8000	3.1
8	Holo		0.1 M Sodium HEPES	7.5	20% w/v PEG 4000 10% v/v isopropanol	2.5

Unusually, all apo datasets collected were at a lower resolution to the crystals prepared in the presence of the mimic. Analysis of the refined holo structures showed that for dataset 5 (Table 5.1), no co-crystallisation was observed as a water molecule could be seen hydrogen-bonding to the active site residue, K122 where the mimic is expected to be bound. The mimic is however clearly aiding crystal packing as the resolution observed for an apo crystal under the same crystallisation condition (dataset 4) was 3 Å, rather than 2.2 Å for the holo structure. In dataset 6, some positive density was observed near to the active site residue E110, however the density was not sufficient to attempt to build in the ligand, perhaps suggesting partial ligand occupancy. In dataset 7, some positive density was observed near residues K122 and F112, but as the resolution of the dataset was poor, the ligand could not be built into the structure with certainty. Dataset 8 showed some promising positive density near E110 and K122, but the resolution was also not sufficient to clearly build the ligand. Therefore, the crystallisation conditions for datasets 6, 7, and 8 were optimised by preparing crystals by the hanging drop method. For the conditions of dataset 8, a four-corner screen of the condition by hanging drop method was performed with and without propan-2-ol present. Other crystals were also sent to the synchrotron, however there was evidence that an increased percentage of cryo-protectant would be required as many crystals were very icy or had dissolved in the loop.

Crystals were only observed in the hanging drop four-corner screen prepared for the conditions of dataset 8, without propan-2-ol present. As there were issues with the cryo-

protectant, glycerol was also tested as a cryo-protectant and the amount used was increased from 20 to 25% in the precipitation conditions. Other crystals grown in the 96-well screens (which appeared after one month) were also sent to the synchrotron. Unfortunately, the resolution observed during screening for many of these sitting-drop crystals, was poor (6 -7 Å), so full datasets were not collected. A range of high-resolution datasets (2-2.2 Å) were collected for the hanging drop crystals prepared, based on the conditions of dataset 8 without propan-2-ol present. Structural refinement was performed for the highest resolution structure (referred to as dataset 9), with precipitation condition of 0.111 M HEPES pH 7.5, 20 % v/v PEG 4000 using 25% glycerol as the cryo-protectant with 10 mM final concentration of the mimic, **190**. A 2 Å dataset was also collected for a crystal prepared via the 96-well screens, in condition 25% w/v PEG 1500, 20% glycerol (referred to as dataset 10). No additional cryo protectant was added. Both crystals had a similar, cubic morphology and were slightly yellow in colour (Figure 5.20). Statistics for both datasets are given in Table 5.2.

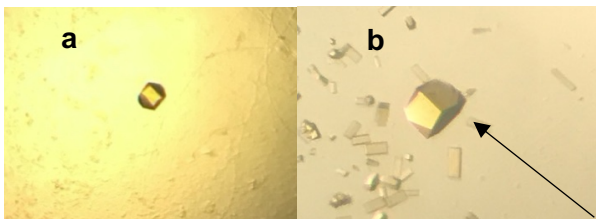


Figure 5.20: **Crystals from which datasets 9 (a) and 10 (b), respectively, were gained.**

Table 5.2: Data collection and refinement statistics of Δ33*Tf*NCS-M97V in complex with reaction intermediate mimic, 190.

	Dataset 9	Dataset 10
Wavelength (Å)	0.976254	0.976254
Space group	$P3_221$	$P3_221$
Unit cell parameters		
a, b, c (Å)	63.66, 63.66, 73.78	62.66, 62.66, 72.65
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Resolution range (Å)	55.13 – 2.22 (2.26-2.22)	54.27 – 2.30 (2.30-2.38)
Total number of observations	95041 (4517)	78587 (6945)
Total number unique	8924 (497)	7696 (719)
Completeness	99.9 (98.8)	99.7 (97.9)
Multiplicity	10.7 (9.1)	10.2 (9.7)
<I/σ(I)>	17.3 (1.1)	26.9 (3.8)
$CC_{1/2}$	0.999 (0.754)	1.000 (0.975)
R_{merge}	0.047 (2.113)	0.035 (0.353)
Solvent content (%)	48	46
Molecule per ASU	1	1
Wilson B factor ($Å^2$)	81.615	69.83
Refinement		
Resolution Range (Å)	55.14 – 1.97 (1.970 – 2.021)	54.27 – 2.30 (2.296 – 2.356)
R_{work}	0.208 (0.242)	0.197 (0.242)
R_{free}	0.295 (0.316)	0.246 (0.302)
Reflection, working	12637	7285
Reflection, free	442	389
Average B factor	70.630	79.715
RMSD bond angle	1.930	1.728
RMSD bond length (Å)	0.0140	0.0128
Ramachandran plot		
Preferred region (%)	89.67	94.87
Allowed region (%)	8.39	4.49
Outliers (%)	1.94	0.64

Both datasets collected were in space group $P3_221$, the same space group observed with the alpha-methyl aldehyde reaction intermediate mimic co-crystallised with wild-type *Tf*NCS. The images for both datasets were integrated using the XDS Dials software programme. Molecular replacement for both datasets was performed using an apo dataset of Δ33*Tf*NCS (wild-type) in the same space group ($P3_221$) as the search model in Phaser.[253] Model building of both the protein and the ligand was performed

using COOT[171] with refinement performed using REFMAC5.[174] TLS refinement was used as there was some anisotropy in the datasets collected. This is based upon the assumption that atoms are part of a rigid body (i.e., a unified group of atoms) of which there is a mean squared displacement. TLS refinement therefore models the anisotropic motions of atoms and can lower the R-factors during refinement.[254] The final R_{work}/R_{free} factors achieved after refinement were 0.21/0.29 and 0.19/0.25 for datasets 9 and 10, respectively. 24 and 50 water molecules were also built into each model (dataset 9 and 10), respectively.

For both datasets positive density was observed in the active site, near to key active site residues K122 and E110. Multiple cycles of model building, and refinement led to the observation of increased density in the active site in both structures, into which the co-crystallised mimic **190** was built. The final ligand placement in both structures, is given in Figure 5.21.

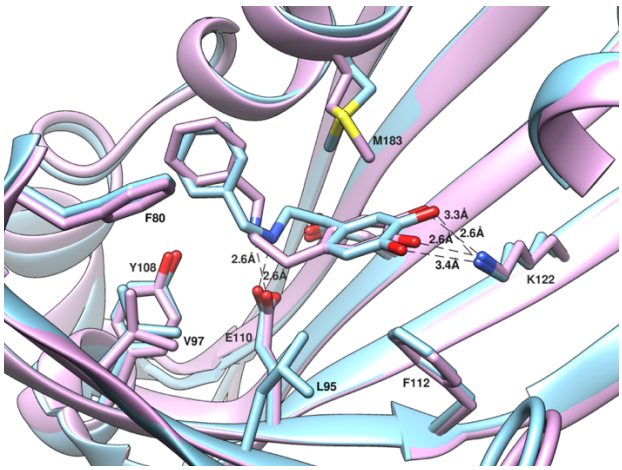


Figure 5.21: **Ligand placement of the refined structures.** Structures are of datasets 9 (blue) and 10 (purple). Image was created using USCF Chimera.[32]

In both structures, the ligand is orientated in a manner consistent with the 'dopamine-first' mechanism. Both catechol hydroxyl groups are hydrogen-bonded to key active site residue, K122. A closer hydrogen-bonding contact is observed for the meta-hydroxyl group (2.6 Å heteroatom to heteroatom distance for both structures), with a weaker interaction observed between K122 and the *para*-hydroxyl of **190** (3.4 and 3.3 Å for

datasets 9 and 10, respectively). These interactions are consistent with previously observed structures of *Tf*NCS with co-crystallised reaction intermediate mimics (PDB: 5NON and 6RP3). It is known that the *meta*-hydroxyl is essential for a productive NCS-mediated reaction. The *para*-hydroxyl is not essential for substrate acceptance, but lower conversions are observed (Table 3.1), likely as an additional hydrogen-bonding interaction assists the deprotonation of the *meta*-hydroxyl. An additional hydrogen-bonding interaction was also observed between the secondary amine of **190** and carboxyl group of E110, also consistent with the 'dopamine-first' mechanism. This residue has been previously shown to be catalytically relevant, in aiding dopamine binding and catalysing iminium ion formation. A variety of different conformations have been observed in different *Tf*NCS structures, depending on the reaction intermediate mimic that is bound.[30,148]

Although the same hydrogen-bonding interactions were observed in both structures (datasets 9 and 10), the ligand placement and the orientation of M183 is different between the two. Despite slightly unclear density in the active site, attempts to force either ligand into the placement of the other were unsuccessful. During analysis of the two ligand structures, it was found that rotation of each of the ligands by 180^{o} results in the overlay of the two. Both ligands are in a pre-cyclisation conformation, with a distance from the aromatic carbon to which attack at the analogous iminium carbon would occur of 3.1 - 3.2 Å. Both are poised to attack on the *Re* face of the *cis* iminium ion, to give a (*S*)-quinone intermediate, which is subsequently deprotonated by E110 to give the product. The NCS active site is very narrow (as shown in Figure 5.22), and it is unlikely that the mimic would be able to rotate in the active site without wider conformational changes to the enzyme's structure. The intramolecular cyclisation of the mimic would require movement of the phenyl group at the 'aldehyde-end' of the molecule, thus requiring widening of the active site and large conformational changes. The NCS

reaction mechanism is known to be highly dynamic, based upon NMR titration experiments.[26]

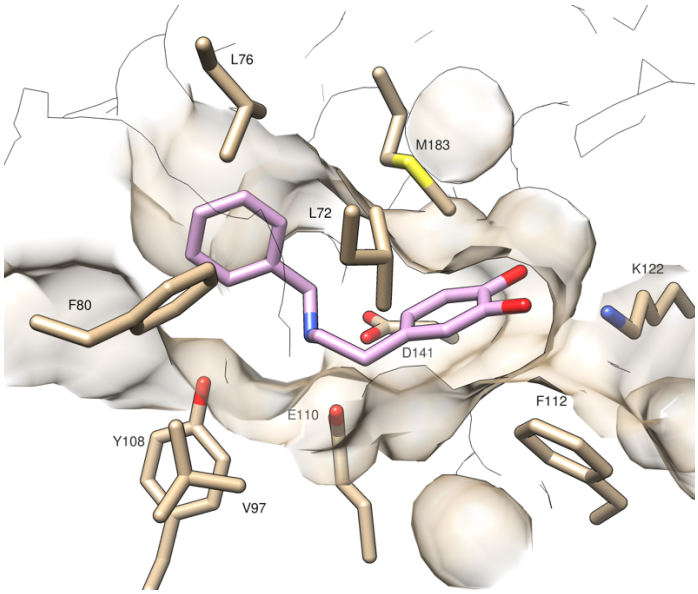


Figure 5.22: **Co-crystallised structure of Δ33*Tf*NCS-M97V with 190 (dataset 10).** A surface model is used to the tight active site pocket in which the ligand residues. Image prepared using Chimera.[32] Figure adapted from Roddan *et al.*[149]

Active site residue M183 is known to be involved in enzymatic activity. Knock-out mutants, M183F/L developed by B. Lichman removed all catalytic activity of the enzyme.[31] In the co-crystallised structure published by Lichman, Sula *et al.* (PDB: 5NON), the terminal methyl group of M183 is oriented away from the ligand whereas in both datasets shown here, the methyl group is directed towards the ligand. Methionine is highly flexible, so it is possible that this is simply a crystallographic artefact.

These ligand conformations are particularly interesting as it is the first time that a pre-cyclisation configuration has been observed with a reaction mimic and it also rationalises the stereochemical outcome of the NCS reaction as only conformations capable of cyclising to give the (*S*)-enantiomer of product are observed in the two different crystals. It is unclear how both ligand conformations occur and the only difference in local protein structure is with M183. Perhaps the ligand can rotate in the active site, modulated by this methionine residue.

A recent paper by Sheng *et al*. rationalises the stereochemistry of the NCS reaction using density functional theory calculations.[33] The reaction pathways to give both the (*S*) and (*R*) enantiomers of product are both favoured, until the final step i.e. the deprotonation of the quinone due to steric clashes of the aldehyde 'R' group with active site residue L72. In the simulations, to orientate the iminium ion before cyclisation occurs, the hydrogen-bonding interaction from the iminium nitrogen to E110 must be cleaved and the iminium nitrogen moves closer to D141. Although no clashes of the phenyl ring with L72 are observed in either crystal structure, these structures only show a single 'snapshot' of the mechanistic process and as discussed, it is likely that conformational changes of the enzyme and mimic would occur. No single point mutants of NCS, altering residue L72 have been generated, so the involvement of this residue in determining the product stereochemistry has not been fully assessed experimentally. Also, the calculations were performed with the natural substrate, 4-HPAA rather than benzaldehyde and an NCS variant was used for these crystallographic studies, rather than the wild-type, and perhaps slightly different residue interactions are occurring.

Residue V97 sits close to the alkyl region of the dopamine end of the putative iminium carbon of the mimic. The mutation from methionine to valine decreases both flexibility, and steric bulk at this position. It is still unclear why this mutant, and the variant M97F increase conversions. The M97V mutant is likely creating extra space in the active site so the reasons for increased conversions are due to faster substrate binding, whereas with the M97F mutant, larger conformational changes across the enzyme may be occurring to create additional space in the active site or there may be a pi-stacking interaction with the benzaldehyde phenyl ring. Regardless, gaining a co-crystallised structure with this mutant would be insightful in understanding why this mutant is still highly productive with these substrates. In an attempt to rationalise the observed activities with M97F, computational docking studies of the iminium ion of reactions

between dopamine (**5**) and benzaldehyde (**139**) with this variant were performed with Autodock Vina, however the results were inconclusive (Figure 5.23). The mimic is orientated with the *para*-hydroxyl group hydrogen-bonded to K122, so likely indicative of a non-productive mode of action, as hypothesised by Lichman *et al*.[30] The ligand is observed in an extended conformation and likely due to narrowing of the enzyme active site entrance / exit. Further structural studies are therefore required to fully understand the improvement in activity of this mutant compared with the wild type.

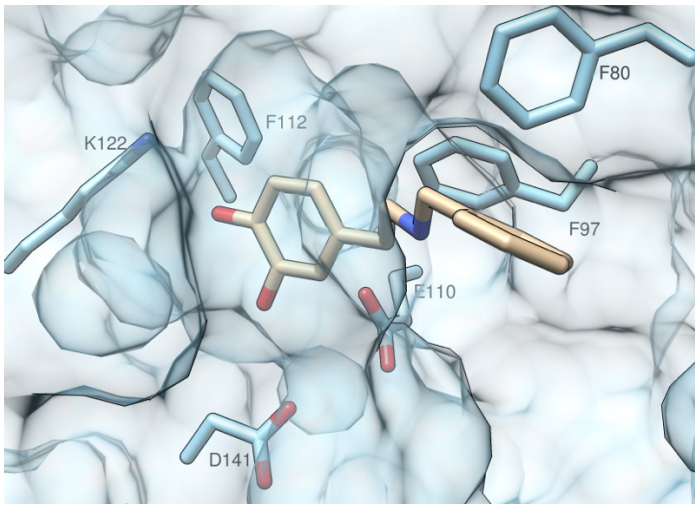


Figure 5.23: **Docking of mimic 190 into the active site of Δ33*Tf*NCS-M97F**. The protein was prepared using NAMD, based upon PDB: 5NON of the wild-type enzyme. Computational docking was performed using Autodock Vina[255] and the figure prepared using USCF Chimera.[32]

To conclude, both co-crystallised structures therefore rationalise benzaldehyde acceptance by *Tf*NCS and confirms the (*S*)-stereochemistry in the THIQ products formed by the enzyme. It is also the first time a non-productive reaction mimic is observed in a pre-cyclisation conformation. The potential role of M183 in determining iminium ion orientation is noted and the M97V substitution is shown to increase space in the active site entrance / exit.

5.3 Conclusions and future work

In conclusion, despite previous reports, a variety of benzaldehydes were shown to be accepted as the carbonyl substrate in *Tf*NCS-mediated reactions with both wild-type and single-point variant enzymes. Previous reactions were performed with low substrate loading (1 mM) and here higher loading is used (10 – 20 mM) due to the low K_d of dopamine binding with *Tf*NCS. The active site mutant, M97V proved particularly promising and could generate a variety of novel (1*S*)-1-aryl THIQ derivatives in high yields and high enantiomeric excesses, particularly with methylated benzaldehyde analogues. This route is advantageous as the chiral THIQ scaffold can be generated in a single step under benign conditions, avoiding the use of high temperatures, high pressures, and toxic reagents. Oxidatively sensitive hydroxyl groups are also tolerated on the THIQ moiety.

High background reactions were observed with a variety of halogenated (fluorinated and chlorinated) benzaldehyde derivatives as substrates but when using high enzyme loading and lowering the reaction pH to 6 the contribution of the spontaneous background reaction was minimised, enzyme activity improved and decreased racemisation at the C-1 position, thus improving the enantiopurity of the resultant THIQ products. A variety of more challenging, hydroxyl, brominated and disubstituted benzaldehyde derivatives were also tested as substrates but were poorly accepted, likely due to steric hinderance or issues of electron deficiency and lack of space in the enzyme active site. Performing NCS reactions with benzaldehydes as substrates is challenging due to their high oxidative sensitivity and reactivity. This was mitigated by using a non-oxidative co-solvent, using purified enzyme, quenching reactions with organic solvent, and storing samples for analysis at low temperatures. High enzyme concentrations are required to rapidly consume the substrates and minimise the contribution of the spontaneous, racemic Pictet-Spengler reaction that occurs.

The products of NCS reactions between dopamine and benzaldehyde or cyclohexanecarboxaldehyde were accepted as substrates by a variety of *O*-methyl transferases. Regioselective methylation of the *meta* hydroxyl was achieved using a SafC enzyme from *Myxococcus xanthus*, which is unusual as with other catecholamines, the methylation is predominantly selective towards the *para*-hydroxyl. Performing regioselective methylation reactions is challenging chemically and the products generated here are pharmaceutically relevant. To generate corresponding *para*-methylated products, 4-methoxytyramine was used as the phenethylamine substrate and the resulting THIQ products from the NCS reactions were obtained in high yields and enantiomeric excesses.

Structural evidence for benzaldehyde acceptance by *Tf*NCS was obtained via a co-crystallised structure of the most productive single point variant, Δ33*Tf*NCS-M97V with a non-productive reaction intermediate analogue (of the reaction between dopamine and benzaldehyde) bound in the active site. The ligand was found to be in a pre-cyclisation conformation, consistent with the generation of products with (*S*)-stereochemistry. The structure also indicated the potential role of M183 in modulating ligand positioning and the single point mutant, M97V removes a flexible residue in the active site entrance/exit, so perhaps this mutation simply provides extra space in the active site for substrates to bind. It is unknown why the active site mutant M97F is also more productive than the wild-type enzyme, but this could be due to pi-stacking effects. Perhaps it is productive as pi-stacking interactions may occur between this residue and the aldehyde phenyl group or that wider conformational changes of the enzyme are occurring.

The mutant M97V proved superior in generating a range of novel (1*S*)-aryl THIQs in high yields and enantiomeric excesses. In future studies, further mutagenesis at this position could be performed, possibly to even smaller sidechains on the amino acid, so

to alanine (M97A) or glycine (M97G). Mutagenesis of neighbouring residues in the active site entrance/exit, L95 and F80, could also be considered. Synthetic derivatisation of the THIQ products generated via Boc protection of the amine proved challenging and poor yields were achieved. Further optimisation of this reaction could be performed to improve conversations and improve the viability of an NCS-mediated step in multi-step syntheses. The expansion of the portfolio of regioselective methyltransferases available will likely lead to increased conversions to give both meta- and para- methoxy products, which would be useful in generating analogues for drug-discovery purposes. The increased conversions using M97F compared with the wild-type have not been fully rationalised and this could be further understood by gaining a co-crystallised structure with this mutant, to see whether wider conformational changes of the enzyme or pi-stacking interactions are observed.

Chapter 6 – *N*-Methyl Phenethylamines as Substrates for Norcoclaurine Synthases

6.1 Introduction

6.1.1 Expanding the phenethylamine substrate scope of *Tf*NCS

The amine substrate scope of *Tf*NCS has been explored to some extent, mostly focussing on altering the catechol moiety. The *meta*-hydroxyl on the phenethylamine is required for turnover however, the *para*-hydroxyl is non-essential, with successful reactions observed with halogen and methoxy- substitutions at this position.[21] Substitutions on the alkyl region of the phenethylamine are also well tolerated with metaraminol and norepinephrine being accepted giving reasonable reaction conversions.[118,256] The stereochemistry of metaraminol (**102**) and norepinephrine (**101**) is unchanged during the Pictet-Spengler reaction, so the use of these catecholamines as the amine substrate in NCS reactions is an excellent way of gaining multiple chiral centres in a molecule in a single reaction step. Example amine substrates tested as substrates by *Tf*NCS are given in Figure 6.1.[21,31]

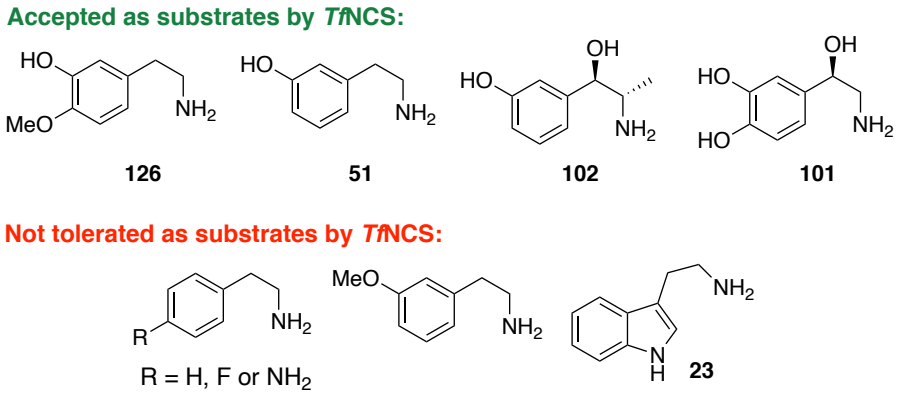


Figure 6.1: The reported phenethylamine substrate scope of *Tf*NCS beyond the natural substrate, dopamine.

There have been no previous reports exploring substitution on the nitrogen atom and expanding the NCS substrate scope is useful to develop synthetic routes towards single enantiomer THIQs. There is a phenomenon known as the 'magic methyl effect' whereby

the addition of a single methyl group to a molecule can dramatically increase the binding affinity to a target.[257] There is therefore interest in trying to find synthetic routes to methylated THIQs to increase pharmacological potency. A variety of *N*-methylated isoquinolines have been shown to exhibit anti-malarial (**89** and **192)** and anti-HIV (**193)** properties *in vitro*.[258] Another *N*-methylated, C-1 substituted THIQ has been shown to have affinity for both D_1 and D_2 dopamine receptors (**194)**, which are targets in drug discovery for the treatment of illnesses such as Parkinson's disease and schizophrenia.[259] (*S*)-Reticuline (**9**), in the biosynthetic pathway to morphine (**1**) also can act as a central nervous system depressant.[260] The structures of various medicinally relevant *N*-methylated isoquinolines and THIQs are given in Figure 6.2.

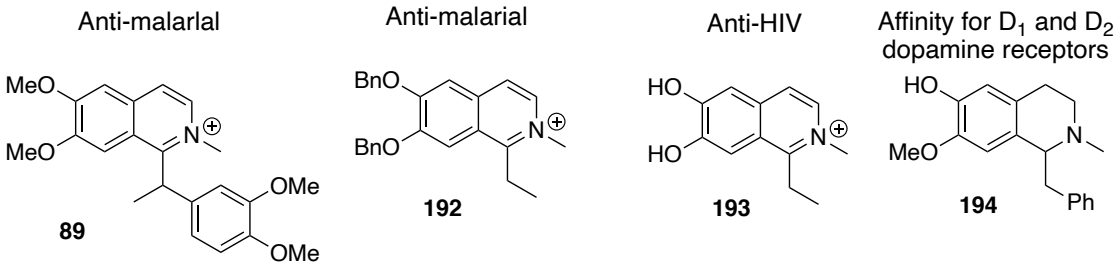


Figure 6.2: Various biologically active *N*-methylated isoquinolines and THIQs.

The synthesis of *N*-methylated THIQs is challenging using traditional organic synthetic methods. Chemical alkylating reagents such as methyl iodide tend to over *N*-methylate, resulting in the formation of quaternary nitrogen centres. These reagents also lack heteroatom selectivity, so synthetic routes often require a protecting group strategy if other heteroatoms are present in the molecule. Alkylating agents are widely known to be carcinogenic and cytotoxic, due to the ability to alkylate DNA.[261] There are some reported biocatalytic routes towards milder and more regioselective *O*- and *N*-methylations. However, the highly selectivity of biocatalysts can also be a hinderance with reported methyltransferases often having a narrow substrate scope.[262] Performing reactions on a larger scale can also be challenging, as the reactions require *S*-adenosyl-L-methionine (SAM) as a co-factor, in stoichiometric quantities. SAM is both unstable and expensive, however recent efforts towards co-factor recycling systems

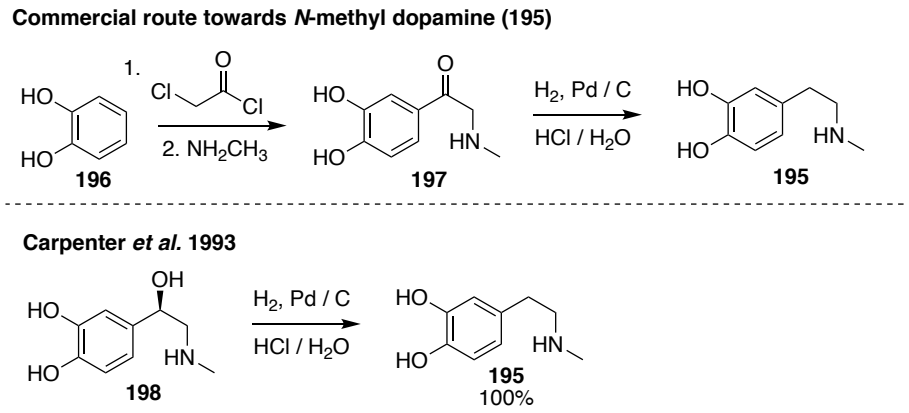
are promising and may allow methyltransferases to become a more economically-viable and benign route to selectively methylated products.[203]

Attempts to synthesise *N*-methylated THIQs using a regioselective *N*-methyltransferase proved unsuccessful when CNMT ((*S*)-coclaurine *N*-methyltransferase isolated from *Coptis japonica*) was tested in the development of chemoenzymatic cascades towards 13-Me-THPBs (Chapter 4). Therefore, to develop an alternative route towards *N*-methylated THIQs, the aim was to screen a range of *N*-methyl-phenethylamines as substrates for NCS. This would hopefully provide a single-step route to a variety of *N*-methylated THIQs without the requirement for a selective methylating reagent after the Pictet-Spengler reaction.

6.2 Results and Discussion

6.2.1 Synthesis of *N*-methyl dopamine

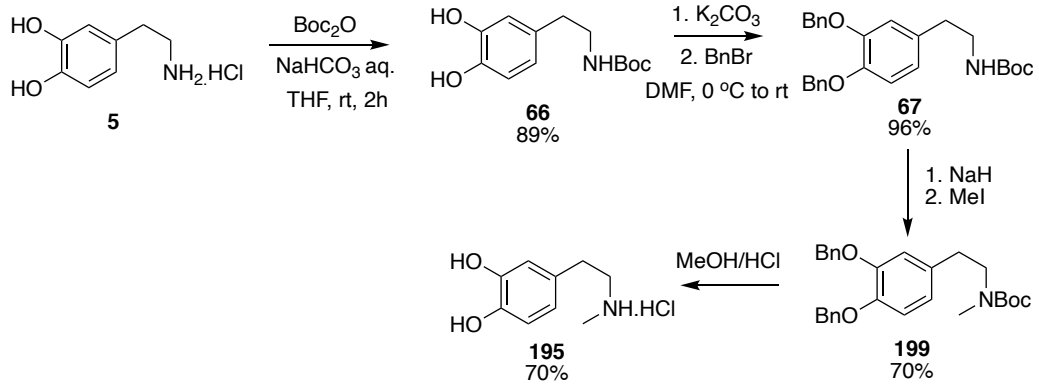
N-methyl dopamine (**195**) was the first selected to screen as it is most structurally similar to dopamine (**5**), the natural phenylethylamine substrate of NCS. Unfortunately, *N*-methyl-dopamine is not commercially available but there are two reported synthetic routes (Scheme 6.1). The commercial route to **195** involves the Friedel-Crafts acylation of catechol (**196**) followed by S_N2 attack with methylamine to give **197**. This was then hydrogenated under acidic conditions at 70 ^oC to give **195**. A simpler route was devised by Carpenter *et al*.[263] giving **195** in a single step starting from epinephrine (**198**). Here, a novel route using a protecting group strategy was developed to enable potential applications in *N*-alkyl analogue synthesis.[263]



Scheme 6.1: Reported synthetic routes towards *N*-methyl dopamine.

Novel synthetic routes towards catecholamines are useful as many are known to act on dopamine receptors and are of interest as treatments for illnesses such as Parkinson's disease. *N*-methyl dopamine (**195**) is also the prodrug of ibopamine, a dopamine agonist, which is used for the treatment of glaucoma[264] and is also under investigation for the treatment of congestive heart failure.[265]

A synthetic route towards *N*-methyl dopamine (**195**) was devised based upon previous work in our group to synthesise secondary amine mimics of the NCS reaction iminium ion intermediate.[30] This involved a protecting group strategy, whereby the catechol hydroxyl groups and amine were protected, allowing only a single methylation at the nitrogen atom, followed by orthogonal deprotection. Most of the reactions are performed at room temperature and literature precedent demonstrated that all reactions are high yielding. Purification is often not necessary at each step as many by-products are volatile and any remaining protecting group starting materials do not hinder further reactions and can also be removed under reduced pressure in the final step.



Scheme 6.2: **Synthetic route performed to give *N*-methyl dopamine.**

Reactions were all high yielding, and the synthesis of *N*-methyl dopamine (**195**) was achieved in four-steps in a 42% overall, isolated yield. The initial Boc protection of dopamine to give **66** was performed using a biphasic system to remove the hydrochloride salt of dopamine and allow for reaction with Boc anhydride. The crude product was obtained with unreacted Boc anhydride as a side-product but was carried through directly to the next step as the excess Boc anhydride could be removed in the final step. Benzyl protection of the catechol was performed using a Williamson ether synthesis followed by recrystallisation to give **67**. Methylation of **67** was challenging, as there were issues with substrate solubility in many organic solvents. Eventually, the reaction was performed in DMF in high yield, however removing the solvent after the reaction was challenging and purification by silica chromatography was required. In order to remove the protecting groups on **199** to give **195**, acid hydrolysis was used followed by the removal of solvents *in vacuo*. This however resulted in a product that was a viscous brown oil, instead of a yellow crystals as reported in the literature.[263] This was rationalised to be due to the presence of a small amount of oxidised product. The oxidation of dopamine (and therefore likely also *N*-methyl dopamine) is *via* a radical polymerisation which produces a conjugated aminochrome product.[137] To isolate *N*-methyl dopamine (**195**), the brown residue was resuspended in acetone. **195** is known to be insoluble in acetone, whereas any organic impurities i.e., the benzyl alcohol side

product, would be soluble. This resulted in the desired final product precipitating as a white solid which could be isolated by filtration in 70% yield.

6.2.2 Initial screening of *N*-methyl phenethylamines as *Tf*NCS substrates

Biocatalytic reactions were performed with **195** and two other *N*-methyl phenethylamines (Figure 6.3), using wild-type Δ33*Tf*NCS as the reaction catalyst. There are two commercially available phenethylamines with an *N*-methyl group and a *meta*-hydroxyl group, which is essential for a productive NCS-mediated reaction; epinephrine (**198**) (also known as adrenaline) and phenylephrine (**200**), which is widely used as a decongestant.[266] Five different aldehydes (**76**, **62**, **78**, **201** and **202,** Figure 6.3) were also screened as the corresponding carbonyl substrate. Phenylacetaldehyde (**76**) was used in place of the natural NCS substrate, 4-HPAA (**6**) as it is commercially available and less prone to oxidation. A variety of linear aldehydes (**62, 78, 201** and **202**) were also screened as they are normally well-accepted as substrates, with conversions comparable to those when using **6**.[31] The two enantiomers of citronellal were also screened, since linear aldehydes were particularly promising as substrates, but unfortunately no conversions were observed with any of the amines (**196**, **198**, or **200**).

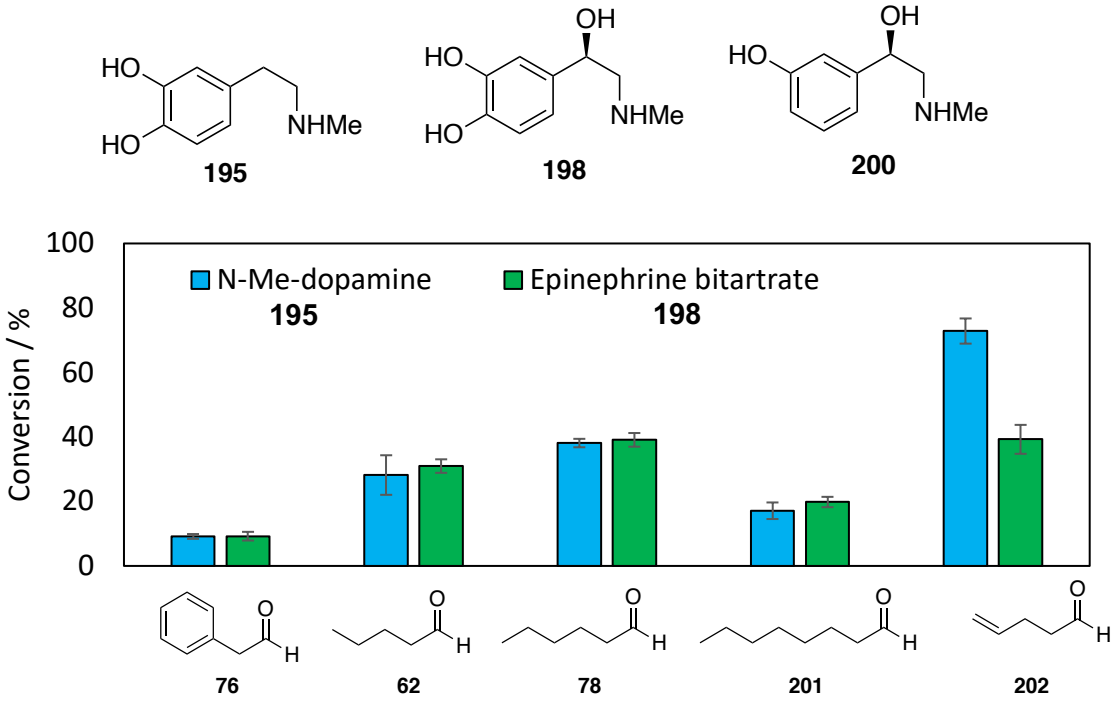


Figure 6.3: ***Tf*NCS reactions with *N*-methyl dopamine hydrochloride or epinephrine bitartrate as the amine substrate**. Aldehydes, phenylacetaldehyde, valeraldehyde, hexanal, octanal or 4-pentenal were used as the carbonyl substrate. *Reaction conditions*: amine (10 mM), aldehyde (20 mM), sodium ascorbate (10 mM), Δ33*Tf*NCS (0.2 mg mL^{-1}) in HEPES buffer (100 mM, pH 7.5) with MeCN (10% *v/v*) at 37 °C, 250 μL scale reactions for 24 h. Yields were determined by a) product formation by analytical HPLC using calibration curves of standards b) amine consumption by analytical HPLC based upon calibration curves of the standard.

Reactions were performed with wild-type *Tf*NCS under standard conditions that were previously developed.[31] For reactions with linear and benzylic aldehydes, and dopamine, 0.1 mg mL^{-1} enzyme has been shown to be sufficient for near complete conversions[31] but instead here, 0.2 mg mL^{-1} enzyme was used as it was suspected that *N*-methyl phenethylamines would be particularly challenging substrates. Trace conversions were observed with **200** as the amine substrate while reasonable conversions were observed using **195** and **198**, as shown in Figure 6.3. The poor acceptance of **200** suggests that the *para*-hydroxyl on the phenyl ring is needed, perhaps to 'anchor' the substrate with the *N*-methyl group into the active site by a catechol hydrogen bonding interaction with K122. Indeed, poorer conversions were also observed when using *meta*-tyramine as a substrate (i.e. dopamine without the *para*-

hydroxyl) compared with using dopamine in earlier work[148] Perhaps the presence of the *N*-methyl group means that there is a weaker interaction with charged residues in the middle of the active site (E110 and D141) to hold the phenethylamine in the enzyme active site. Therefore, another hydrogen bonding interaction (such as *para*-hydroxyl of the phenethylamine to the amine of K122) might be required for successful turnover. It was considered that this is likely due to lack of space in the active site or weaker interactions with active site residues crucial for turnover, as the amine of the phenethylamine is hindered by the methyl group. Interestingly, the best conversions were observed with 4-pentenal (**202**) as the aldehyde substrate. With **195** as the amine, around a 50% increase in conversion was observed simply by introducing the alkene group. This could be due to favourable π-stacking interactions with the side chains of residues F80 and Y108 in the NCS active site entrance/exit. Regardless, using **202** as the aldehyde substrate is useful as the alkene in the THIQ product could be used as a metathesis handle to further diversify the products.

Characterisation of the products generated by NMR spectroscopy indicated that two conformers were present, probably due to inversion of the nitrogen lone pair leading to two different methyl group chemical environments. High temperature (60 ^{o}C) ^{1}H-NMR analysis of the product of the reaction between *N*-methyl dopamine (**195**) and **62** was performed to confirm this. In the product, two singlets at around 6.6 ppm are expected, corresponding to the two aromatic protons. This is observed at 60 ^{o}C but at room temperature, two conformers are present, resulting in another pair of two aromatic singlets, as shown in Figure 6.4.

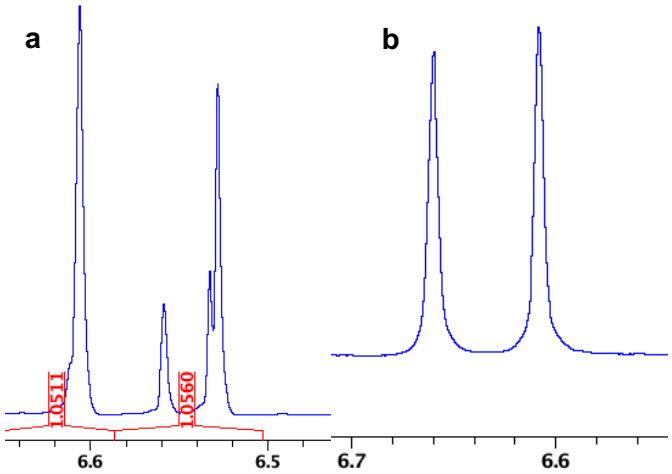


Figure 6.4: **Aromatic region of ^{1}H-NMR spectrum of an *N*-methylated THIQ generated by the NCS mediated reaction between 195 and 62 a**) Room temperature and **b**) high temperature (60 ^{o}C) ^{1}H-NMR spectra (6.7-6.5 ppm).

RP-HPLC and LCMS analysis of the products generated in reactions with epinephrine bitartrate suggested that two diastereomers of product were present. From literature precedent, the major enantiomer of product formed from an NCS reactions has (*S*)-stereochemistry at the C-1 position,[267] so it can be assumed that the major diastereomer of product formed is (1*S*,4*R*). This therefore suggested that a small amount of the (1*R*,4*R*) product was generated as it was highly unlikely that the mild reaction conditions used would racemise the chiral centre at C-4.

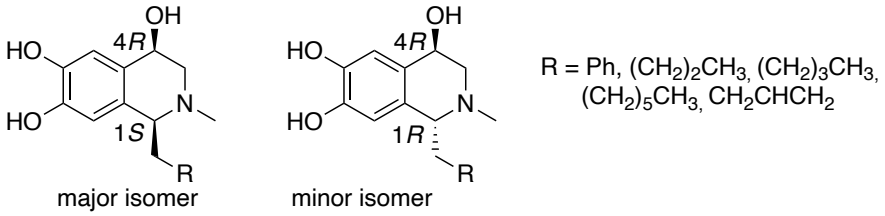


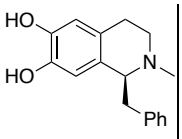
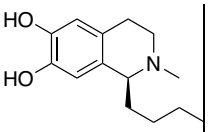
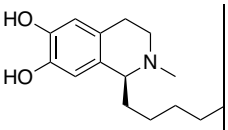
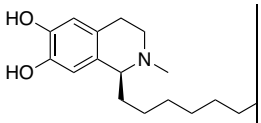
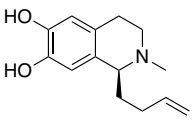
Figure 6.5 The two isomers of product generated by NCS reactions using epinephrine (198) as the phenethylamine substrate.

To probe whether the (1*R*)-product was formed, chiral HPLC analysis was performed with the products of NCS reactions with **195**, to give **203**-**207**, as only one chiral centre is generated and for NCS catalysed reactions between dopamine **5** and a variety of

aldehydes, the THIQ products are usually formed in high e.e. (often >95%),[35] Chiral HPLC separations were achieved using AD-H columns for all THIQs generated, except the THIQ generated from the reaction with hexanal (**205**) which was surprising considering that the THIQs generated from reactions with pentanal and octanal (**206** and **207**) were separated.

The chiral HPLC results confirmed that with an *N*-methyl group on the phenethylamine, the high stereoselectivity mediated by the NCS reaction was lost (Table 6.1). From mechanistic studies with NCS, it is known that the interaction of the nitrogen on the amine substrate with the two charged residues, E110 and D141 is crucial for mediating iminium ion formation and may therefore from this work be implicated in determining the stereoselectivity of the reaction during cyclisation. Perhaps the introduction of a methyl group on nitrogen results in a lack of appropriate interaction with these residues, thus leading to a loss of stereoselectivity in the product.[28]

Table 6.1: **Enantiomeric excesses of the products of NCS mediated reactions between *N*-methyl dopamine and five different aldehydes**. Reactions were performed using 0.2 mg mL^{-1} of purified *Tf*NCS under standard conditions. nd = e.e. not determined. The major isomer of product is shown.

				
203	**204**	**205**	**206**	**207**
8% e.e.	29% e.e.	nd	20% e.e.	10% e.e.

In an attempt to improve the yields and selectivities, reactions were performed using the same conditions, but at increased enzyme concentrations and also when lowering the pH of the reaction to 6. Indeed, lower pH had been crucial to gaining high enantiomeric excesses in the (1*S*)-aryl-THIQs generated between dopamine and numerous halogenated benzaldehydes (Chapter 5), where the reduction in pH

appeared to minimise the background, racemic Pictet-Spengler reaction and improve enzyme activity. Reactions for screening were performed using the same five aldehydes (**76**, **62**, **78**, **201** and **202**, Figure 6.3) and 3-phenylpropanal (**203**) was also utilised as **202** had been particularly promising as a substrate (Figure 6.6).

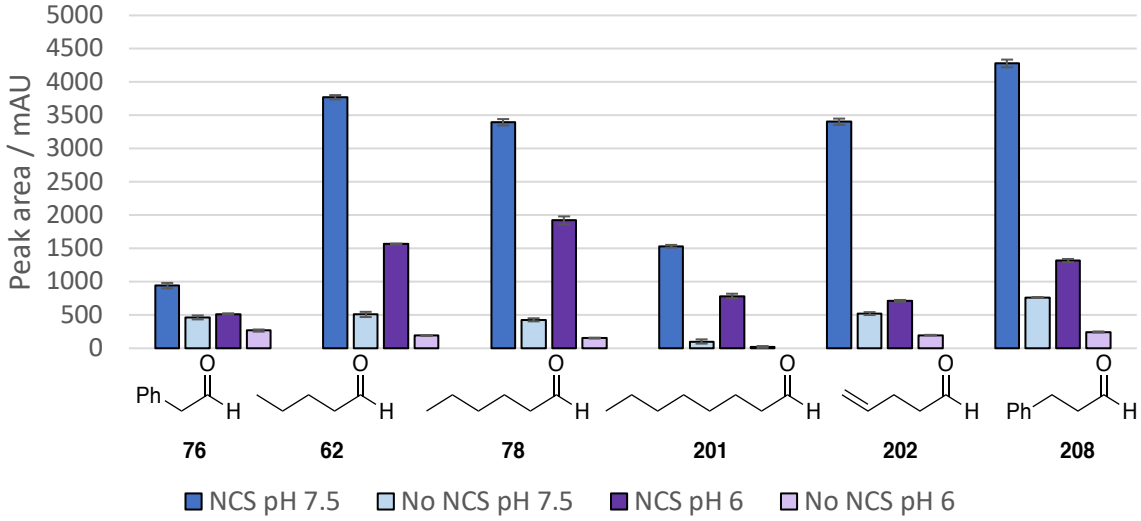


Figure 6.6: **WT-*Tf*NCS catalysed reactions between *N*-methyl dopamine and a range of aldehydes.** Reactions were performed with 0.5 mg mL^{-1} *Tf*NCS at either pH 6 or pH 7.5.

Yields and enantiomeric excesses of products are given in Table 6.2. The product generated for reactions with 3-phenylpropanal (**203**) was not isolated (but instead was confirmed by LCMS), so yields and e.e.s are not given. Unfortunately, both yields and selectivity were negatively affected by performing reactions at a lower pH. Fairly high background reactions were also observed, likely due to the electron donating methyl group on the nitrogen increasing nucleophilicity and imine formation, meaning that spontaneous, racemic background Pictet-Spengler reactions were more likely. However, the background reactions are not significant enough to result in such poor reaction selectivities, meaning that the low e.e.s are associated with the NCS sample, as may also be the case with the benzaldehydes as discussed in Chapter 5. This could be further investigated by using a knockout variant as a negative control, perhaps based upon one of the double point variant NCSs as discussed in Chapter 2.

Table 6.2: **Yields and enantiomeric excesses of products generated by *Tf*NCS reactions using 0.5 mg mL^{-1} enzyme at either pH 6 or pH 7**. Yields were determined based upon calibration curves of isolated products. Enantiomeric excesses were determined by chiral HPLC (method 5).

THIQ	pH 6		pH 7.5	
	Yield / %	e.e. / %	Yield / %	e.e. / %
203	19	-	36	-
204	42	36	>99	55
205	43	-	77	-
206	21	32	41	46
207	24	0	>99	27

6.2.3 Screening of NCS variants

Previous biocatalytic studies have shown that certain active site mutants of norcoclaurine synthase can improve conversions with certain substrates. For example, the mutant M97V can improve conversions with α-substituted aldehydes as substrates,[148] while the mutant A97I can improve yields with ketones.[36] With this knowledge, a variety of different NCS mutants (L76V, A79I, A79F, F80L, M97V, M97L, M97F and Y108F) were screened in the reactions between *N*-methyl dopamine and phenylacetaldehyde (Figure 6.7a) and epinephrine bitartrate and valeraldehyde (Figure 6.7b). Reactions were performed using clarified NCS cell lysates as the reaction catalyst.

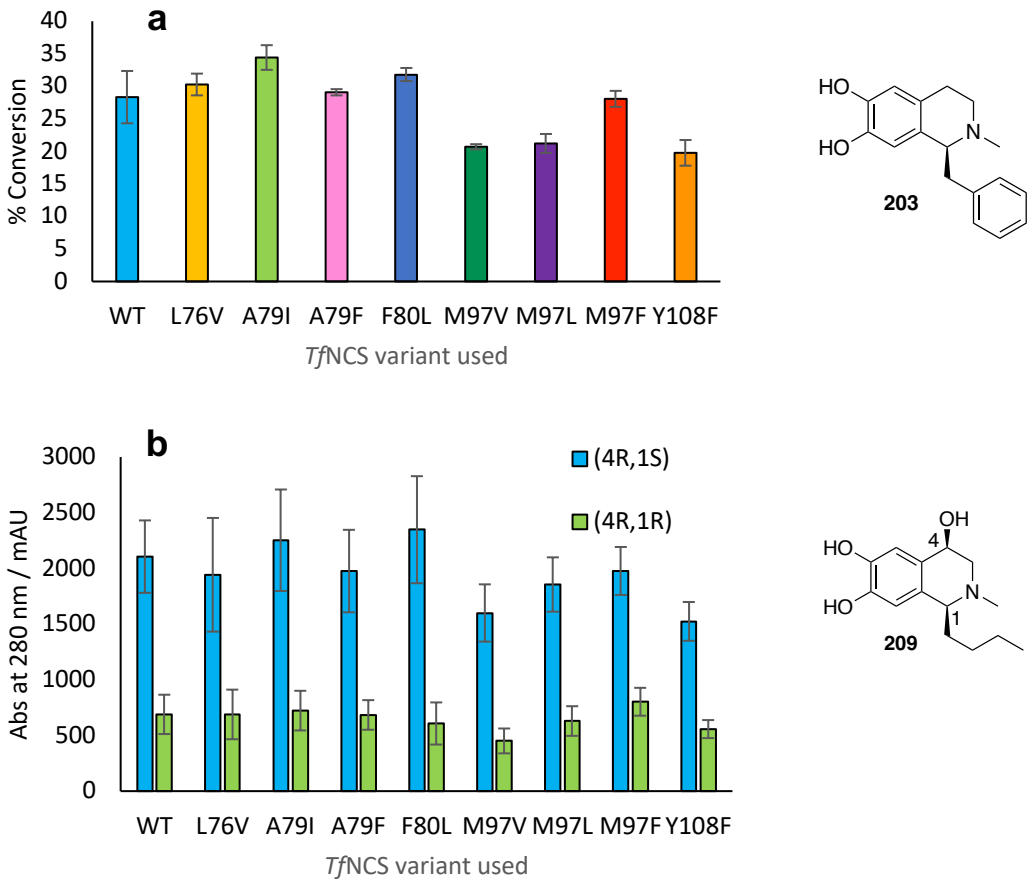


Figure 6.7: Screening a library of NCS single point variants with *N*-methyl phenethylamines as substrates. Structures of the major isomer of product are given for each. **a**. HPLC conversions by amine consumption to give **203. b**. Product peak areas by RP-HPLC of the reaction between epinephrine bitartrate (**198**) and pentanal (**62**), to give **209**. *Reaction conditions*: amine (10 mM), aldehyde (20 mM), sodium ascorbate (10 mM), Δ33*Tf*NCS (0.2 mg mL^{-1}) in HEPES buffer (100 mM, pH 7.5) with MeCN (10% *v/v*) at 37 °C, 250 mL scale reactions for 24 h.

For the formation of **204**, improved conversions were observed with the mutant A79I. This was also the case with many other NCS substrates such as ketones and aliphatic aldehydes[31,36] and is rationalised as being due to increased hydrophobicity of the active site and so enhanced substrate binding.[28] For reactions with epinephrine bitartrate (**198**), a single enantiomer of epinephrine was used and it is unlikely that this chiral centre will be racemised during the reaction. Therefore, the presence of two diastereomers would be due to altered stereochemistry at the C-1 position. The two diastereomers of product (4*R*,1*S*) or (4*R*,1*R*) are easily separable by RP-HPLC. For all

the mutants tested, there was little change in the diastereomeric ratios of the two products, however no mutants of the residues at E110 and D141 were tested. From mechanistic studies, these two residues are likely to orient the iminium ion during cyclisation and thus removing the hydrogen-bonding interaction with these two residues may result in a less stereoselective reaction.[28]

6.2.4 Phosphate reactions to generate racemic products

To generate racemic standards for chiral HPLC analysis and to provide a comparison to the enzymatic reaction, reactions were performed between **195** and aldehydes **76**, **62**, **78**, **201** and **202**, but using inorganic phosphate instead of NCS as the reaction catalyst.[15] For all substrates tested, complete amine consumption was observed, reflecting the high activation of the amine due to electron donation by the *N*-methyl group. Conversions could not be accurately quantified by product formation calibration curves as two regioisomers of product were formed. For reactions with *N*-methyl dopamine, the two regioisomers were formed in a 3:1 ratio of *para*:*ortho* products by NMR spectroscopic analysis. The two regioisomers were separated by preparative-HPLC (method 9) and the *para* isomer was used as a racemic standard for chiral HPLC analyses.

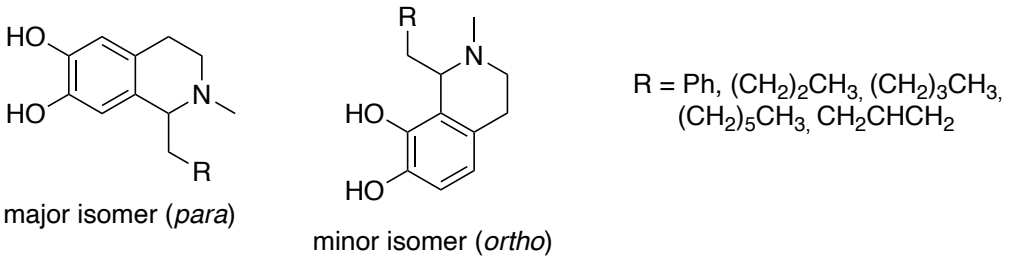
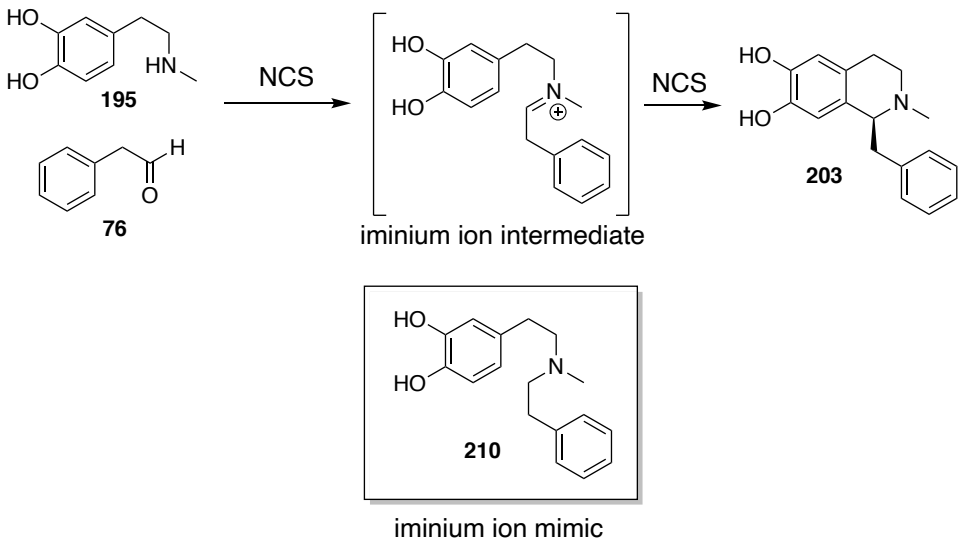


Figure 6.8: The two regioisomers of product generated by a phosphate-mediated reaction between *N*-methyl dopamine and a variety of aldehydes.

6.2.5 Structural studies to understand the reaction selectivity

Protein crystallography has been used successfully to rationalise NCS substrate acceptance by the co-crystallisation of the enzyme with non-productive reaction

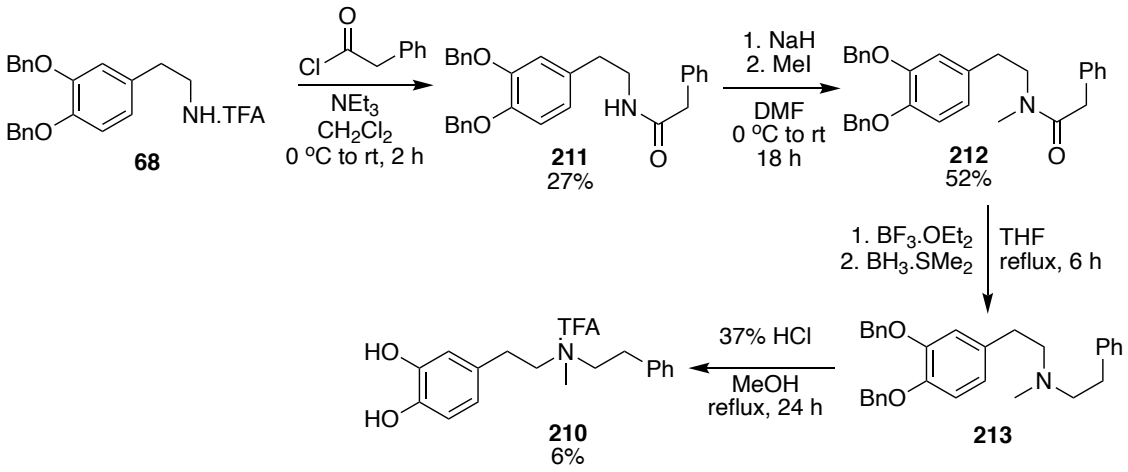
intermediate mimics.[148,149] Therefore, to try to understand the poor conversions and stereoselectivities observed in products generated by NCS reactions with *N*-methyl phenethylamines, a reaction intermediate analogue of the reaction between *N*-methyl dopamine (**195**) and phenylacetaldehyde (**76**) was synthesised. The rationale behind the mimic design (i.e., comparison to the iminium ion intermediate of the reaction) is given in Scheme 6.3 Previous attempts to co-crystallise reaction mimics based on reactions with linear aldehydes had proved unsuccessful due to a combination of poor difference density observed in the active site and multiple ligand conformations. Therefore, despite aldehydes **62**, **78**, **201** and **202** being preferred as the carbonyl substrate partner for reactions with **195, 76** was chosen as the aldehyde part of the mimic. It was hoped that the phenyl ring would provide additional electron density to aid observation in the enzyme active site and that the mimic would be less flexible, to avoid previous issues of multiple ligand conformations.



Scheme 6.3: **Design of a *N*-methyl dopamine reaction intermediate analogue, based upon the iminium ion intermediate of the NCS reaction.**

The mimic, **210** was synthesised in four steps from **68** (Scheme 6.4). Amide coupling between **68** and phenylacetyl chloride proceeded to give **211** in 27% isolated yield. The *N*-methyl group was then incorporated by reaction with iodomethane in the presence of base. Some di-methylation was observed so the mono-methylated product, **212** was

isolated by column chromatography in 52% yield. Rotamers were observed by NMR spectroscopic analysis of **208** and this was resolved by increasing the temperature at which the spectra were collected (60 °C vs. rt). The amide was reduced to the secondary amine (**213**) using borane and taken through directly to the final reaction step, benzyl deprotection by heating at reflux under acidic conditions. The final product, **210** was isolated as the TFA salt in 6% yield by preparative HPLC. A poor yield was obtained due to incomplete benzyl deprotection and would likely be improved by increasing the reaction time.



Scheme 6.4 Synthetic route towards an *N*-methyl dopamine reaction intermediate analogue for co-crystallisation experiments with *Tf*NCS.

The mimic (**210**) was prepared at 200 mM in DMSO and added to Δ33*Tf*NCS (12.6 mg mL^{-1} in 20 mM TRIS, 50 mM NaCl, pH 7.5) to give 5% *v/v* mimic in the protein sample. Despite extensive screening, crystals were only observed under three conditions which is a stark contrast from screening with other mimics where many crystals were gained. From crystals sent to the synchrotron, only apo protein structures were observed. The lack of co-crystallisation likely reflects the poor acceptance of *N*-methyl dopamine (**195**). The mimic was also prepared in water or DMSO. and soaking of some apo crystals was attempted; however, the crystals were destroyed upon addition of the mimic. Soaking with water or DMSO alone did not cause the crystals to dissolve, indicating that this was due to the mimic disrupting the crystal lattice. Although enzymes can retain activity

in the crystal lattice[29] and *Tf*NCS is known to have a highly dynamic mechanism,[26] crystal destruction was not observed when soaking other reaction mimics which were based upon more productive enzymatic reactions.

Alternatively, to gain structural insights into *N*-methyl phenethylamine acceptance, computational docking studies were performed using AutoDock Vina. The iminium ion of the reaction between *N*-methyl dopamine (**195**) and phenylacetaldehyde (**76**) was used for docking into the active site as this would help gain information on the orientation of the substrates before cyclisation. Chain A of *Δ*33*Tf*NCS (PDB: 5NON) was used, with the ligand removed. No electrostatic interactions are observed between the nitrogen of the iminium ion and any of the heteroatoms of key charged residues in the mechanism (E110, D141 and Y108), as shown in Figure 6.9. This is likely to due to steric hindrance from the *N*-methyl group.

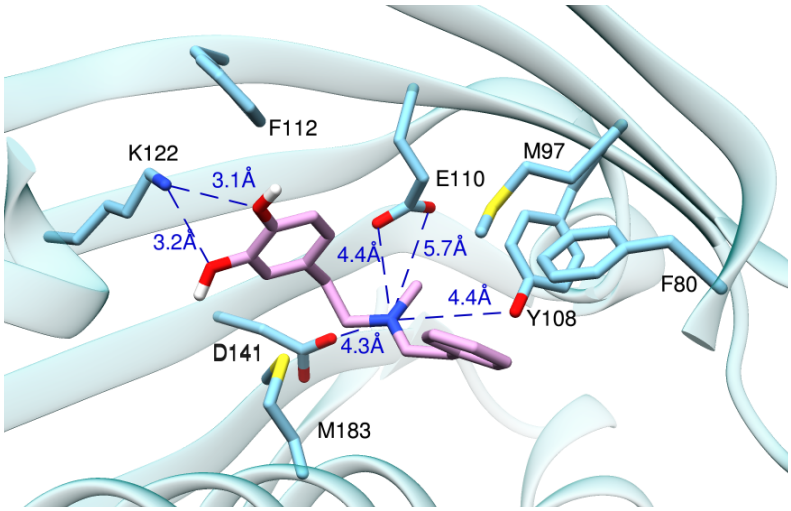


Figure 6.9: **Placement of iminium ion reaction intermediate of the reaction between *N*-methyl dopamine and phenylacetaldehyde in the active site of *Tf*NCS**. Residues 65-79 were removed to aid visualisation and the figure was made using USCF Chimera.[32] Hydrogen bonding distances are shown by blue dotted lines.

6.3 Conclusions and future work

In summary, work described in this chapter demonstrates that *Tf*NCS can accept *N*-methylated catecholamines as substrates to give a variety of THIQ products in high

regioselectivities, however the stereoselectivity and yields of reactions are poor. Conversely, with inorganic phosphate as the reaction catalyst, yields are high but there is a lack of stereocontrol. Computational docking studies suggested that this could be due to a lack of interaction between the charged residues, E110 and D141, and the iminium ion intermediate. A variety of NCS mutants were tested with the aim of improving conversions compared with the wild type. Conversions were improved with the active site mutant, A97I, likely due to increased hydrophobicity of the active site. Although substrates of this type were accepted by *Tf*NCS, a more favourable route to single enantiomer *N*-methylated THIQs would likely be *via* a biocatalytic cascade involving a first step by NCS to generate the THIQ scaffold in a regio- and stereoselective manner, followed by selective mono *N*-methylation using an appropriate selective *N*-methyltransferase.

Chapter 7: Conclusions and Future Work

7.1 Conclusions

This work has resulted in the extension of the known substrate scope of *Tf*NCS, particularly towards α-methyl-substituted aldehydes and the benzaldehydes. This provides novel stereoselective routes towards pharmaceutically relevant THIQs. Further functionalisation of the NCS-generated THIQs was performed by regioselective methyltransferase enzymes and using chemical Pictet-Spengler reactions. Crystallographic investigations of *Tf*NCS by the co-crystallisation of reaction intermediate mimics were successfully used to rationalise the acceptance of reactions with α-substituted aldehydes.

In Chapter 2, biocatalytic routes towards (*R*)-THIQs are presented. A range of enzymatic variants were explored; *Tf*NCS variants, NCSs from *N. nucifera* and the salsolinol synthase enzyme. Although no changes in stereoselectivity were observed, NCS proved to be highly resilient, retaining stereoselectivity despite significant active-site mutagenesis.

Based upon previous reports of α-substituted aldehydes generally being poorly accepted by *Tf*NCS, Chapter 3 describes the use of α-methyl-substituted aldehydes as substrates. Indeed, a range of α-methyl-substituted aldehydes were shown to be well-accepted as substrates. When a racemic aldehyde was used, the (*R*)-enantiomer was preferentially accepted leading to the regio- and stereoselective syntheses of (1*S*,1'*R*)-1'-Me-THIQs. Active site variants improved activities compared to the wild-type enzyme and crystallographic studies with a co-crystallised reaction intermediate mimic helped to rationalise these improved activities.

Chapter 4 describes the development of chemoenzymatic cascades towards 13-Me-THPB and 13-Me-PB alkaloids like the biologically active alkaloids isolated in *Corydalis*

plants. The cascades involved regioselective catechol *O*-methyltransferase and chemical Pictet-Spengler reactions.

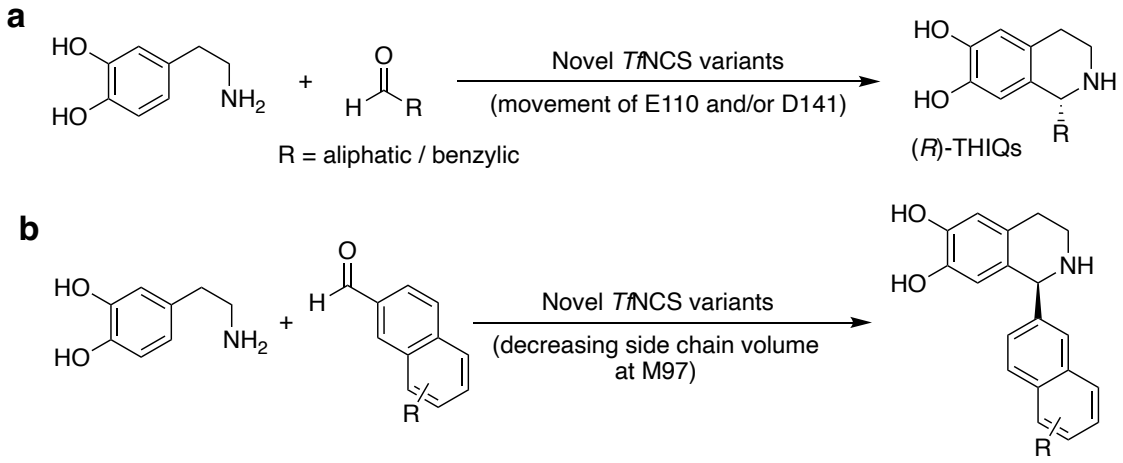
The improved acceptance of α-methyl-substituted aldehydes, particularly by the *Tf*NCS variant, M97V led to the screening of benzaldehydes as substrates, as described in chapter 5. After optimisation of the reaction conditions, numerous (1*S*)-aryl-THIQs were generated in high yields and e.e.s. A co-crystallised structure of this M97V variant helped rationalise the observed activities. Methylations with catechol *O*-methyltransferases were also performed and the selectivities of the enzymes explored.

N-Methyl-phenethylamines were investigated as *Tf*NCS substrates, and the results are presented in Chapter 6. Work focused on *N*-methyl-dopamine which was shown to be accepted by the enzyme. The stereoselectivity of the reaction was lost although products were still generated in a regioselective manner.

7.2 Future Work

NCS has proved to be an excellent biocatalyst for generating (*S*)-THIQs in high yields and with high regio- and stereoselective control. Future work could focus on further redesign of the enzyme, either by rational design or directed evolution. Particularly, these could build on the efforts made in Chapter 2 to develop an NCS that can synthesise (*R*)-THIQs. Changing the positions of the two, charged active site residues, E110 and D141 to alter the conformation that the iminium ion reaction intermediate adopts may lead to a change in stereoselectivity. Generating new constructs of the salsolinol synthase and the sacred lotus NCSs may also result in active enzymes that can generate (*R*)-THIQs. The reticuline epimerase enzyme (Scheme 1.12) converts (*S*)-reticuline to (*R*)-reticuline. Significant enzyme engineering efforts of this enzyme may also provide a route towards (*R*)-THIQs (as shown in Scheme 7.1a). The variant M97V has already proved promising in improving the acceptance of α-substituted

aldehydes by *Tf*NCS. Therefore, mutations focusing on decreasing steric bulk and side chain flexibility at this position (e.g. M97A and M79G) may allow access to more sterically challenging aldehydes such as 2-napthaldehyde and analogues as shown in Scheme 7.1b). Deacetylipecoside synthase (DIS) is another known Pictet-Spenglerase that has been shown to perform an *R*-selective reaction (Figure 1.5).[92] The sequence and substrate scope of the enzyme is unknown, although is hypothesized to have similarities with strictosidine synthase. This enzyme may therefore provide a stereoselective route to (R)-THIQs.



Scheme 7.1: Ideas for future enzyme engineering efforts to generate novel, enantiopure THIQs.

Although crystallographic studies were used successfully in this work to rationalise observed NCS activities, gaining high resolution structures with the ligand bound in a single conformation with high occupancy was challenging. The highly dynamic nature of the NCS mechanism meant that few mechanistic insights could be drawn from NMR studies performed in 2008.[26] However, based on what is now known about structure and mechanistically relevant residues of NCS, further understanding of the dynamics of the mechanism by NMR studies may be possible, providing a useful guide for enzyme engineering efforts.

Numerous novel alkaloids were also generated in this work, many based upon compounds that are known to be biologically active. Therefore, selected compounds

have been sent to collaborators to assess their activities. Based upon activities with similar compounds, malaria, zika virus, anti-bacterial and dopamine receptor inhibitory activities are to be explored.

Chapter 8: Experimental

8.1 Enzyme details

Table 8.1: Plasmids used for this work. All are stored as plasmid DNA and as glycerol stocks in *E. coli* BL21(DE3). SDM = site directed mutagenesis. GS = gene synthesis (purchased from [a]DNA 2.0, [b]Eurofins or Life Technologies with codon optimisation for *E. coli* expression and NdeI/XhoI restriction sites). PCR = gene amplified by colony PCR from NEB 5-alpha *E. coli*.[209] The sequence and preparation of all Δ29*Tf*NCSs and wild type Δ33*Tf*NCS has been previously described in the PhD thesis of B. Lichman and published.[30,31] The subcloning and sequences of two methyltransferases (*Rn*COMT and *Mx*SafC)[209,211] and recycling system enzymes (*Ec*MAT, *Ec*MTAN)[203] has been reported. *Cj*6OMT and *Cj*CNMT have also been reported.[197] pQR codes of these enzymes are not given as they were provided by other groups.

pQR	Plasmid	Gene	Mutation(s)	His-tag terminal	Source	Method
1046	pJ411	Δ29*Tf*NCS	Wild type	C	B. Lichman	GS[a]
1045	pJ411	Δ29*Tf*NCS	L76V	C	B. Lichman	GS[a]
1047	pJ411	Δ29*Tf*NCS	A79I	C	B. Lichman	GS[a]
1833	pJ411	Δ29*Tf*NCS	F80L	C	B. Lichman	GS[a]
1841	pJ411	Δ29*Tf*NCS	M97V	C	B. Lichman	SDM
1840	pJ411	Δ29*Tf*NCS	M97L	C	B. Lichman	SDM
1839	pJ411	Δ29*Tf*NCS	M97F	C	B. Lichman	SDM
1834	pJ411	Δ29*Tf*NCS	Y108F	C	B. Lichman	GS[a]
1846	pJ411	Δ29*Tf*NCS	K122L A69K	C	B. Lichman	SDM
1847	pJ411	Δ29*Tf*NCS	K122L L72K	C	B. Lichman	SDM
1848	pJ411	Δ29*Tf*NCS	K122L I85K	C	B. Lichman	SDM
1849	pJ411	Δ29*Tf*NCS	K122L F112K	C	B. Lichman	SDM
1850	pJ411	Δ29*Tf*NCS	K122L L95K	C	B. Lichman	SDM
1851	pJ411	Δ29*Tf*NCS	K122L M126K	C	B. Lichman	SDM
1852	pJ411	Δ29*Tf*NCS	K122L M183K	C	B. Lichman	SDM
1853	pJ411	Δ29*Tf*NCS	K122F F112K	C	B. Lichman	SDM
1856	pD451-SR	Δ33*Tf*NCS	Wild type	N-TEV	B. Lichman	GS[a]
1637	pET29a(+)	Δ33*Tf*NCS	M97V	N-TEV	J. Ward	GS[b]
1648	pET29a(+)	Δ33*Tf*NCS	L72A	N-TEV	Y. Wang	GS[b]
1649	pET29a(+)	Δ33*Tf*NCS	L72V	N-TEV	Y. Wang	GS[b]
1650	pET29a(+)	Δ33*Tf*NCS	Y108S	N-TEV	Y. Wang	GS[b]
1651	pET29a(+)	Δ33*Tf*NCS	Y108W	N-TEV	Y. Wang	GS[b]
1665	pET29a(+)	Δ33*Tf*NCS	P179A	N-TEV	R. Roddan	GS[b]
1667	pET29a(+)	Δ33*Tf*NCS	P179G	N-TEV	R. Roddan	GS[b]
1669	pET29a(+)	Δ33*Tf*NCS	P179S	N-TEV	R. Roddan	GS[b]
1671	pET29a(+)	Δ33*Tf*NCS	F80W	N-TEV	R. Roddan	GS[b]
1673	pET29a(+)	Δ33*Tf*NCS	F80Y	N-TEV	R. Roddan	GS[b]
2339	pET29a(+)	*Nn*NCS5	Wild type	N-TEV	J. Ward	GS[b]
2340	pET29a(+)	*Nn*NCS7	Wild type	N-TEV	J. Ward	GS[b]
-	pET28a(+)	*Rn*COMT	Wild type	N	J. Andexer	GS[c]
-	pET20b(+)	*Mx*SafC	Wild type	C	J. Andexer	-
-	pET28a-SacB-SapI	*Cj*6OMT	Wild type	C	F. Subrizi	GS[b]
-	pET28a(+)	*Cj*CNMT	Wild type	C	F. Subrizi	-
-	pET28a(+)	*Ec*MAT	Wild type	N	J. Andexer	PCR
-	pET28a(+)	*Ec*MTAN	Wild type	N	J. Andexer	PCR
1663	pEX-A128	SS v1	Wild type	C	J. Ward	GS[b]
1664	pEX-A128	SS v2	Wild type	C	J. Ward	GS[b]
1665	pEX-A128	SS v3	Wild type	C	J. Ward	GS[b]

8.2 Molecular Biology Methods

8.2.1 Transformations into *E. coli*

Transformations were performed with either Stellar™ Competent Cells (Takaka) or BL21(DE3) competent cells (NEB). On ice, DNA (2 μL) was added to competent cells (50 μL) and incubated for 30 min. Cells were then heat shocked (42 °C, 30 s) and incubated on ice (5 min). 500 μL of Super Optimal Broth with Catabolite repression (SOC media) was added, and the cells were incubated at 37 °C (250 rpm, 1 h). The cells were then plated onto lysogeny broth (LB) agar, containing the appropriate antibiotic (kanamycin at 50 μg mL^{-1}) and incubated overnight at 37 °C.

8.2.2 Plasmid isolation

LB media (10 mL) with appropriate antibiotic (kanamycin at 50 μg mL^{-1}) was inoculated with a single colony of *E. coli* containing the desired plasmid. Cultures were incubated overnight at 37 °C, 250 rpm. Cells were isolated by centrifugation (5,000 rcf, 5 min, 4 °C) and DNA isolated by mini-prep kit (Sigma Aldrich, GenElute™Plasmid Miniprep Kit) following the manufacturer's instructions. DNA was eluted into nuclease free water and concentrations were measured using a NanoDrop Spectrophotometer.

8.2.3 Restriction digests

1-2 μg DNA, 5 μL 10X CutSmart Buffer (NEB), 1 μL NdeI restriction enzyme (NEB) and 1 μL XhoI restriction enzyme (NEB) were prepared in nuclease free water with a final volume of 50 μL. Digests were performed for 1 h at 37 °C then DNA fragments separated by agarose gel electrophoresis.

8.2.4 DNA gel electrophoresis

1% agarose gels in 1 x TBE buffer (0.089 M TRIS HCl, 0.089 M boric acid, 0.002 M EDTA) were prepared with SYBR Safe DNA gel stain (0.03% *v/v*). Purple loading dye

(6X, 10 μL) was added to each 50 μL restriction digest reaction. Gels were run at 110 V, 40 min in 1 x TBE buffer.

8.2.5 Gel extraction

The desired fragments of DNA were cut from the agarose gels and DNA isolated using a Qiagen MinElute® Gel Extraction Kit.

8.2.6 Ligations

Digested and isolated vector and insert DNA fragments were combined at a molar ratio of vector:insert of 1:7 with 20 ng of vector DNA. Reactions were performed using Instant Sticky-end Ligase Master Mix (NEB), following the manufacturers instruction then placed on ice and transformed into Stellar™ Competent Cells (Takaka).

8.3 Protein preparation

8.3.1 Protein expression

Expression of *Tf*NCS (wild-type and variants of both Δ29 and Δ33 constructs), *Nn*NCS5 and salsolinol synthases was performed using the following protocol. All plasmids were used to transform *E. coli* BL21(DE3). 10 mL LB media was inoculated with a single colony and grown at 37 °C for 18 h. 500 mL LB media was inoculated with 1% v/v of overnight culture and grown at 37 °C for 2 h. IPTG (0.5 mM final concentration) was added, and the protein was incubated for 18 h at 25 ° C. Cell pellets were isolated by centrifugation and stored at -20 °C until further purification or lysis for enzymatic reactions.

This expression protocol was adapted for the other enzymes investigated. For the expression of *Rn*COMT, *Mx*SafC, 6OMT, *Ec*MAT and *Ec*MTAN, the same expression protocol was used, except for 0.2 mM final concentration of IPTG was used instead. For the expression of 6OMT, the protein was instead incubated at 18 °C for 36 h. For

the expression of the *Nn*NCS7, induction was performed with 0.5 mM IPTG (final concentration) with cultures incubated at 18 ^{o}C for 24 h. For the expression of *Mx*SafC, where incubation was performed at 16 ^{o}C for 36 h.

8.3.2 Lysate preparation

To prepare lysates for enzymatic reactions, the pellet was resuspended in buffer (4% v/v of final culture volume, 50 mM HEPES, pH 7.5) and cells were lysed by sonication (10 s ON, 10 s OFF, 10 x). The resulting suspension was centrifuged (10 min, 6000 x g, 4 ^{o}C) and the supernatant used directly for enzymatic reactions or stored at -20 ^{o}C. The total protein concentration in the lysate was determined by Bradford assay and SDS-PAGE analysis combined with quantification using ImageJ, was used to determine the percentage of *Tf*NCS in the lysate.

8.3.3 Protein purification

For purification of all *Tf*NCSs, *Nn*NCSs, *Mx*SafC and *Ec*MAT, the cell pellet was resuspended in lysis buffer (100 mM HEPES, 20 mM imidazole, 100 mM NaCl, pH 7.5, 15% v/v culture volume) with a small amount of DNAse1. The cells were lysed by sonication (3 x [3 min ON, 3 min OFF]) and centrifuged (15,000 x g, 45 min, 4 ^{o}C). The resulting supernatant was removed and filtered (0.45 µm). A 5 mL His-trap HP column was equilibrated with lysis buffer and the lysate loaded onto the column at 1 mL min^{-1}. The column was washed with lysis buffer to remove any unbound protein (6 column volumes), followed by washings with a stepwise gradient of increasing imidazole concentrations by combining lysis buffer and elution buffer (100 mM HEPES, 500 mM imidazole, 100 mM NaCl, pH 7.5). The column was washed with 8% elution buffer for 6 column volumes, 16% elution buffer for 6 column volumes followed by 100% elution buffer. All washings were performed at a flow rate of 1 mL min^{-1}. Fractions were analysed by SDS-PAGE and fractions containing pure protein were combined and dialysed into 20 mM Tris, 50 mM NaCl, pH 7.5 buffer. The protein sample was

centrifuged and concentrated to approximately 10 mg mL^{-1}. Aliquots of protein were stored at -80 °C until use for enzymatic reactions or in the case of Δ33*Tf*NCS-WT or M97V, taken through to the next step of purification for removal of the hexahistidine tag.

For removal of the hexahistidine tag of Δ33*Tf*NCS-WT and M97V, TEV protease was added to the pooled fractions and the sample dialysed overnight (20 mM Tris, 50 mM NaCl, pH 7.5). The protein sample was then centrifuged (15,000 x g, 20 min, 4 °C) and loaded onto a 5 mL His-trap HP column. The supernatant was passed through the column. The column was then washed with varying amounts of elution buffer (20 mM TRIS, 50 mM NaCl, 500 mM imidazole, pH 7.5) mixed with the dialysis buffer. Fractions containing Δ33*Tf*NCS were pooled and further purified by gel filtration on a Superdex S75 column. Fractions within the major UV absorbance peak at 280 nm were pooled, concentrated (to approximately 10 mg mL^{-1}) and stored at -80 °C.

For purification of *Ec*MTAN, the pellet was resuspended in buffer (4% *v/v* final culture volume, 50 mM HEPES, pH 7.5) and cells were lysed by sonication (10 s ON, 10 s OFF, 10 x). The resulting suspension was centrifuged (10 min, 6000 x g, 4 °C) and filtered (0.45 μM). A Ni-NTA gravity column (7 mL volume) was washed with lysis buffer (100 mM HEPES, 20 mM imidazole, 100 mM NaCl, pH 7.5) and the lysate supernatant loaded onto the column. Any unbound protein was removed by washing with lysate buffer (10 column volumes), then 40 mM imidazole buffer (100 mM HEPES, 40 mM imidazole, 100 mM NaCl, pH 7.5, 10 column volumes). The pure protein was eluted with 500 mM imidazole buffer (100 mM HEPES, 500 mM imidazole, 100 mM NaCl, pH 7.5). 2 mL fractions were collected and analysed with Bradford reagent to test for presence of the protein. Fractions containing high concentrations of protein were pooled and buffer exchanged using a PD-10 column, into 50 mM HEPES, 20 mM NaCl, pH 7.5. The pure protein sample pooled based upon SDS-PAGE analysis of protein fractions and stored in aliquots at -80 °C.

For the purification of salsolinol synthase, the pellet was resuspended in buffer (5% v/v of final culture volume, 50 mM TRIS, pH 7.4) and cells were lysed by sonication (10 s ON, 10 s OFF, 10 x). 1 M perchloric acid was added (5 x resuspension volume) and the solution centrifuged (5000 x g, 10 min, 4 ^{o}C). The supernatant was removed, filtered (0.22 μM) and analysed by SDS-PAGE.[147]

8.3.4 Protein visualisation and concentration determination

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed using Bolt™4-12% Bis-Tris Plus gels with 1 x MES running buffer was used to run the gels at 200 V for 20 min. Gels were stained using InstantBlue™.

Anti-hexahistidine tag western blot analysis was used to visualise proteins with poor expression (*Nn*NCSs and salsolinol synthases). SDS-PAGE was used to separate proteins in the sample, which were transferred to a nitrocellulose membrane using iBlot PVDF transfer stacks (ThermoFischer). The membrane was incubated for 1 h at rt in blocking solution (1 x PBS + 5% non-fat dry milk (Marvel)). The blocking solution was removed, and the membrane incubated with 3 x 50 mL washing solution (1 x PBS + 0.05% TWEEN® 20) every 5 min. 20 mL antibody solution (1 x PBS + 1% non-fat dry milk + 0.05% TWEEN + 10 μL anti-polyHistidine Peroxidase conjugate antibody in 0.01 M Na_2PO_4, 1% BSA, 0.01% thimasol) was added and the membrane incubated for 2 h at rt with rocking. The antibody solution was discarded, and the membrane washed with washing solution (3 x 50 mL, replaced every 5 min). The membrane was developed using SiGMA*FAST* 3,3'-diaminobenzidine following the manufacturer's instructions. The reaction was stopped by removal of the development solution and incubation of the membrane in dH_2O.

Purified protein concentrations were determined using a NanoDrop® spectrophotometer, measuring the absorbance of light at 280 nm of 2 μL of sample. Extinction co-efficients were calculated using ExPASy ProtParam[268] and protein concentrations estimated using the Beer-Lambert Law (Equation 1).

$$c = \frac{A_{280}}{\varepsilon_{280} \times \ell} \times M_W \qquad \text{Equation 1}$$

Protein percentage in cell lysates was determined using ImageJ (NIH) analysis of SDS-PAGE. Total protein concentrations in clarified cell lysate were determined using Bradford protein assays. Percentage of *Tf*NCS in the lysate was determined by ImageJ analysis of the SDS-PAGE gel.

8.4 Biotransformations

8.4.1 NCS catalysed reactions

A solution of amine (final reaction concentration of 10 mM) and sodium ascorbate (final reaction concentration of 10 mM to prevent catecholamine and aldehyde oxidation) was prepared in HEPES buffer (50 mM, pH 7.5). A solution of aldehyde (200 mM in MeCN or DMSO) was prepared. The solutions were mixed to give 10 mM amine, 10 mM sodium ascorbate, 20 mM aldehyde and *Tf*NCS (at approximately 10 mg mL^{-1} in 20 mM Tris, 50 mM NaCl, pH 7.5) was added (at a final concentration of 0.1, 0.2, 0.5 or 1 mg mL^{-1} as described for each reaction) in HEPES buffer (50 mM pH 7.5) with 10% *v/v* MeCN or DMSO. Reactions were shaken or stirred at 37 °C for 3, 18 or 24 h as described. Analytical scale reactions were prepared to give a final reaction volume of 100 μL or 200 μL. Preparative scale reactions were performed on a 10 mL scale.

For reactions using *meta*-tyramine as the amine substrate, sodium ascorbate was not added as there was no catechol-containing moiety in the reaction mixture. To quench

or workup the reactions, different conditions were used depending on the use of the sample, as detailed in Section 8.4.3.

8.4.2 Methyltransferase reactions

The lyophilised NCS reaction (workup method 5) was resuspended in buffer (250 mM HEPES, 200 mM $MgCl_2$, 2 M KCl, pH 7.5) and water (1:4 ratio) in an equal volume to the NCS reaction sample. ATP (100 mM in H_2O, final concentration 10 mM), L-methionine (100 mM in H_2O, final concentration 10 mM), *Ec*MAT (9 mg mL^{-1} purified, 0.4 mg mL^{-1} final concentration), *Ec*MTAN (2.5 mg mL^{-1} purified, 0.025 mg mL^{-1} final concentration) and methyltransferase (*Rn*COMT: 10% *v/v* lysate, 6OMT: 10% *v/v* lysate, *Mx*SafC: 6 mg mL^{-1} purified, 0.6 mg mL^{-1} final concentration) were added. Reaction volumes were adjusted by addition of HEPES buffer (50 mM, pH 7.5). Reactions were performed at 37 °C, 500 rpm (200 μL scale) or 150 rpm (10-20 mL scale).

8.4.3 Workup methods for product isolation or analytical analysis

Workup method 1 (for RP-HPLC analysis and chiral HPLC analysis with methods 3-4): Reactions were quenched by the addition of an equal volume of MeCN. Samples were centrifuged (16,000 x g, 10 min, 4 °C), then the supernatant diluted with an equal volume of water and analysed immediately by analytical RP-HPLC or stored at -20 °C.

Workup method 2 (for chiral HPLC analysis with methods 5-6): The reaction was saturated by addition of $NaHCO_3$ and an equal volume of EtOAc to the reaction volume was added. The samples were vortexed (30 s) and centrifuged (16,000 x g, 5 min, 4 °C). The organic layer was removed and left to evaporate until dryness (*ca.* 48 h). The sample was resuspended in 80:20:0.1 *n*-hexane:ethanol:diethylamine(1.5 x initial reaction volume) and injected for HPLC analysis.

Workup method 3 (to prepare for preparative HPLC purification or crude NMR analysis): The reaction was saturated by addition of $NaHCO_3$ and extracted with ethyl acetate. The organic phases were dried with Na_2SO_4 and concentrated under reduced pressure.

Workup method 4 (to prepare for preparative HPLC purification or crude NMR analysis): The reaction mixture was extracted with ethyl acetate (3 x reaction volume). The organic phases were combined, dried with Na_2SO_4 and concentrated under reduced pressure.

Workup method 5 (to prepare for preparative HPLC purification): Reactions were quenched by the addition of an equal volume of MeCN. Samples were centrifuged (16,000 x g, 10 min, 4 ^oC) and the supernatant removed and lyophilised.

8.5 Analytical methods

8.5.1 Achiral analytical HPLC analysis

Analysis of samples was performed using an Agilent 1260 Infinity liquid chromatography system comprising of a G1329B autosampler, G1311C quaternary pump, 1260 G1316A column oven and 1260 G1314F variable wavelength detector. The system was equipped with a HiChrom ACE C18 column (250 mm x 4.6 mm). Acetonitrile (MeCN) in water (0.1% v/v TFA) was used as the mobile phase.

HPLC method 1: The gradient used was: 10% MeCN for 1 min, a linear gradient to 70% MeCN over 5 min, 100% MeCN for 30 s, followed by 10% MeCN for 3.5 min. A flow rate of 1 mL min^{-1} was used. 10 μL of sample was injected. UV absorbance was measured at 280 nm.

HPLC method 2: The gradient used was: 10% MeCN for 1 min, a linear gradient to 70% MeCN over 15 min, 100% MeCN for 30 s, followed by 10% MeCN for 3.5 min. A flow rate of 1 mL min^{-1} was used. 10 μL of sample was injected. UV absorbance was measured at 280 nm.

8.5.2 Chiral analytical HPLC analysis

HPLC method 3: A Supelco Astec Chirobiotic™ T2 column (250 mm x 4.6 mm) was used with an isocratic MeOH (0.1% AcOH, 0.2% TEA) mobile phase at 1 mL min^{-1}. UV absorbance at 280 nm was measured. The column temperature used was 30 ^{o}C.

HPLC method 4: An Agilent InfinityLab Poroshell 120 Chiral-V column (150 mm x 4.6 mm, 2.7 microns) was used with an isocratic MeOH (0.2 wt% ammonium formate) mobile phase at 1 mL min^{-1}. UV absorbance at 280 nm was measured. The column temperature used was 30 ^{o}C.

HPLC method 5: A Diacel Chiralpak® AD-H column (250 mm x 4.6 mm, 5 μM) was used with an isocratic *n*-hexane:ethanol:diethylamine (80:20:0.01) mobile phase at 1 mL min^{-1}. UV absorbance at 280 nm was measured. The column temperature used was 30 ^{o}C.

HPLC method 6: A Diacel Chiralpal® OD column (250 mm x 4.6 mm, 5 μM) was used with an isocratic *n*-hexane:isopropanol:diethylamine (90:10:0.1) mobile phase at 1 mL min^{-1}. UV absorbance was measured at 280 nm. The column temperature used was 20 ^{o}C.[250]

HPLC method 7: A Diacel Chiralpal® OD column (250 mm x 4.6 mm, 5 μM) was used with an isocratic *n*-hexane:isopropanol:diethylamine (95:5:0.1) mobile phase

at 1 mL min^{-1}. UV absorbance was measured at 280 nm. The column temperature used was 20 °C.[250]

8.5.3 Semi-preparative HPLC

Reversed-phase, semi-preparative HPLC was performed on a Dionex 580 HPLC machine with a PDA-100 photodiode array detector, P580 Pump, and a model ASI-100 automated sample injector. The column used was a Phenomenex Onyx C18 (100 x 10 mm). Water (0.1% TFA) and acetonitrile (0.1% TFA) were used as the mobile phase using a linear gradient of 5-95% acetonitrile over 36 min. A flow rate of 2 mL min^{-1} was used with detection at 254 nm.

8.5.4 Preparative HPLC

Purification was performed using Agilent 1200 Infinity series liquid chromatography system equipped with G1361A prep pump, G2260A autosampler, G1364A fraction collector and G7165 multiple wavelength detector. A Supelco Discovery®BIO Wide Pore C18-10 column (25 cm x 21.2 mm, 10 μM) using MeCN and water as the mobile phase (both solvents contained 0.1% *v/v* TFA) as the mobile phase. UV absorbance was measured at 280 nm.

HPLC method 8: 20-minute gradient 5-95% MeCN in water.

HPLC method 9: 20-minute gradient 25-95% MeCN in water.

HPLC method 10: 20-minute gradient 35-80% MeCN in water.

HPLC method 11: 30-minute gradient 5-95% MeCN in water.

HPLC method 12: 20-minute gradient 25-30% MeCN in water.

8.5.5 GC analysis

Samples were prepared at 10 mM in ethyl acetate and analysed using an Agilent 7820A Gas Chromatograph. 5 μL of sample was injected onto a chiral column (Beta DEX 225, fused silica capillary column 30 m x 0.25 mm x 0.25 μm). A step gradient was used: 1,

40 °C for 1 min. 2, Ramp to 100 °C at 40 °C min^{-1}. 3, Retained temperature at 100 °C for 8 min. 4, Increase temperature to 200 °C at a gradient of 40 °C min^{-1}. 5, 200 °C for 1 min. This method was devised by Helena Philpott.

8.6 General chemical methods

Silica column chromatography was performed using Geduran® Si 60 silica (43-60 μM). Thin layer chromatography analysis was performed using plates with a silica gel matrix on an aluminium support. Ultraviolet light (254 nm) was used to visualise thin layer chromatography plates. All reagents were obtained from commercial sources (Sigma Aldrich, Fischer, ChemCruz, Alfa Aesar) unless otherwise specified.

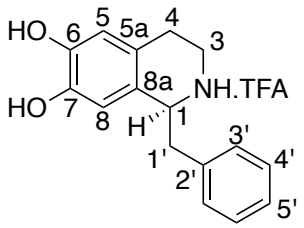
^{1}H and ^{13}C NMR spectra were obtained using Bruker Advance III 600 MHz and 700 MHz spectrometers (as specified in the characterization data). Chemical shifts specified are relative to trimethylsilane (set at 0 ppm) and referenced to the residual, protonated NMR solvent. Coupling constants in ^{1}H-NMR spectra (*J*) are given in Hertz (Hz) and described as singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m). Data processing was performed using the Bruker Topspin 4.0.8 software package.

Mass spectrometry data was obtained using a Waters Aquity UPLC-MS system (LRMS [MS ES+]) or a Waters LCT Premier XE ESI Q-TOF mass spectrometer (HRMS). Infrared (IR) data was obtained using a Bruker Alpha Platinum-ATR machine. Optical rotations (α_D) were determined using a Bellingham and Stanley ADP 430 polarimeter.

8.7 Characterization of THIQs generated from enzymatic reactions

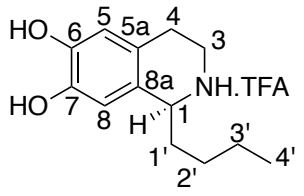
8.7.1 Products from NCS reactions between dopamine and benzylic or linear aldehydes

(*S*)-1-Benzyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (77)[269]



An NCS-mediated reaction was performed using wild-type *Tf*NCS (0.1 mg mL^{-1} final concentration) on a 10 mL, 10 mM scale, followed by workup method 5 and purification by preparative HPLC (method 8) to give the product as the TFA salt (16 mg, 45%). ^{1}H-NMR (600 MHz; D_2O) δ 7.55 - 7.46 (3H, m, 5'-H and 3'-H), 7.45 – 7.41 (2H, m, 4'-H), 6.87 (1H, s, 8-H), 6.80 (1H, s, 5-H), 3.64 – 3.58 (2H, m, 3-*H*H or 1'-*H*H), 3.41 – 3.36 (1H, m, 3-*H*H or 1'-*H*H), 3.24 – 3.17 (1H, m, 3-*H*H or 1'-*H*H), 3.14 - 3.00 (2H, m, 4-H_2); ^{13}C-NMR (176 MHz; D_2O) δ 144.8 (C-6 or C-7), 143.5 (C-6 or C-7), 135.7, 130.1, 129.8, 128.4, 124.4, 123.9, 116.4 (C-5), 114.5 (C-8), 56.7 (C-1), 39.8 (C-3 or C-1'), 39.3 (C-3 or C-1'), 24.6 (C-4); *m/z* [ES$^+$] 256 ([M+H]$^+$, 100%); retention time (achiral analytical HPLC method 1) = 5.7 min.

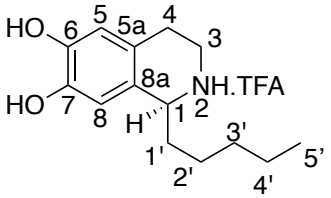
(*S*)-1-Butyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (64)[269]



An NCS-mediated reaction was performed using wild-type *Tf*NCS (0.1 mg mL^{-1} final concentration) on a 10 mL, 10 mM scale, followed by workup method 5 and purification by preparative HPLC (method 8) to give the product as the TFA salt (2 mg, 6.2%). ^{1}H-

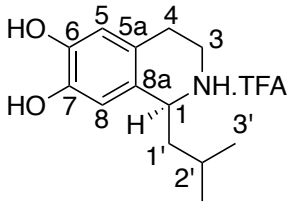
NMR (600 MHz; CD_3OD) δ 6.65 (1H, s, 8-H), 6.60 (1H, s, 5-H), 4.37 – 4.34 (1H, m, 1-H), 3.53 – 3.48 (1H, m, 3-*H*H), 3.35 – 3.27 (1H, m, 3-H*H*), 3.02 – 2.95 (1H, m, 4-*H*H), 2.90 (1H, dt, *J* = 17.1, 5.9 Hz, 4-H*H*), 2.08 – 2.00 (1H, m, 1'-H), 1.92 – 1.83 (1H, m, 1'-H), 1.52 – 1.40 (4H, m, 2'-H and 3'-H), 0.99 (3H, t, *J* = 7.1 Hz, 4'-H); ^{13}C-NMR (151 MHz; CD_3OD) δ 146.5 (C-6 or C-7), 145.8 (C-6 or C-7), 124.2 (C-8a), 123.4 (C-5a), 116.2 (C-5), 113.8 (C-8), 61.5 (C-1), 41.0 (C-3), 34.8 (C-1'), 28.6 (C-4), 25.7 (C-2'), 20.9 (C-3'), 14.2 (C-4'); *m/z* [ES+] 222 ([M+H]$^+$, 100%).

(*S*)-1-Pentyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (79)[269]



An NCS-mediated reaction was performed using wild-type *Tf*NCS (0.1 mg mL^{-1} final concentration) on a 10 mL, 10 mM scale, followed by workup method 5 and purification by preparative HPLC (method 8) to give the product as the TFA salt (17 mg, 53%). ^{1}H-NMR (600 MHz; CD_3OD) δ 6.64 (1H, s, 8-H), 6.60 (1H, s, 5-H), 4.63 (1H, s, 2-H), 4.34 – 4.30 (1H, m, 1-H), 3.53 – 3.48 (1H, m, 3-*H*H), 3.34 – 3.26 (1H, m, 3-H*H*), 3.02 – 2.94 (1H, m, 4-*H*H), 2.90 (1H, dt, *J* = 17.2, 5.9 Hz, 4-H*H*), 2.07 – 1.98 (1H, m, 1'-*H*H), 1.90 – 1.82 (1H, m, 1'-H*H*), 1.55 – 1.34 (6H, m, 2'-H, 3'-H and 4'-H), 0.95 (3H, t, *J* = 6.9 Hz, 5'-H), ^{1}H-NMR (151 MHz; CD_3OD) δ 146.7 (C-6 or C-7), 145.9 (C-6 or C-7), 124.3 (C-8a), 123.6 (C-5a), 116.2 (C-5), 113.8 (C-8), 56.7 (C-1), 41.0 (C-3), 35.1 (C-1'), 32.7 (C-3'), 26.1 (C-4 or C-2'), 25.7 (C-4 or C-2'), 23.5 (C-4'), 14.3 (C-5'); *m/z* [ES+] 236 ([M+H]$^+$, 100%); retention time (achiral analytical HPLC method 1) = 6.0 min.

(*S*)-1-Isobutyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (65)[269]

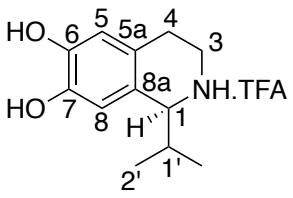


An NCS-mediated reaction was performed using wild-type *Tf*NCS (0.1 mg mL^{-1} final concentration) on a 10 mL, 10 mM scale, followed by workup method 6 and purification by preparative HPLC (method 8) to give the product as the TFA salt (3 mg, 10%). ^{1}H-NMR (600 MHz; CD_3OD) δ 6.62 (1H, s, 5-H or 8-H), 6.61 (1H, s, 5-H or 8-H), 4.42 – 4.37 (1H, m, 1-H), 3.52 – 3.46 (1H, m, 3-*H*H), 3.35 – 3.29 (1H, m, 3-H*H*), 3.00 – 2.88 (2H, m, 4-H_2), 1.86 – 1.72 (3H, m, 1'-H and 2'-H), 1.07 (6H, t, *J* = 6.9 Hz, 3'-H); ^{13}C-NMR (150 MHz; CD_3OD) δ 146.6 (C-6 or C-7), 145.8 (C-6 or C-7), 124.7 (C-8a), 123.7 (C-5a), 116.2 (C-5), 113.9 (C-8), 54.4 (C-1), 49.9 (C-1'), 40.2 (C-3), 25.7 (C-4), 23.5 (C-2'), 22.0 (C-3'); *m/z* [ES+] 222 ($[M+H]^+$, 100%).

8.7.2 Products from NCS reactions with α-methyl and α-ethyl substituted aldehydes

NCS-mediated reactions for product isolation were performed using wild-type *Tf*NCS (0.1 mg mL^{-1} final concentration of purified enzyme) on a 10 mL, 10 mM scale followed by workup method 6 and purification by preparative HPLC (method 1). Products were isolated as the TFA salt.

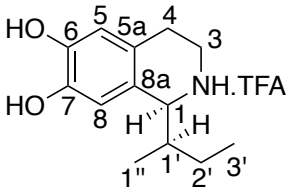
(*S*)-1-Isopropyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (93)[148]



The product was isolated as a white solid (11 mg, 34%). $[\alpha]_D^{25}$ -0.103 (c = 0.69, MeOH); υ_{max}/ cm^{-1} (thin film): 3300-2500, 1673, 1613, 1534; ^{1}H-NMR (700 MHz; D_2O) δ 6.83

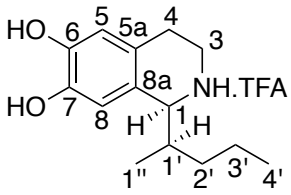
(1H, s, 8-H), 6.80 (1H, s, 5-H), 4.41 (1H, d, *J* = 5.8 Hz, 1-H), 3.64 – 3.59 (1H, m, 3-*H*H), 3.34 – 3.28 (1H, m, 3-H*H*), 3.07 – 3.00 (1H, m, 4-*H*H), 2.96 – 2.89 (1H, m, 4-H*H*), 2.56 – 2.48 (1H, m, 1'-H), 1.14 (3H, d, *J* = 8.4 Hz, 2'-H), 0.88 (3H, d, *J* = 8.5 Hz, 2'-H); ^{13}C-NMR (176 MHz; D_2O) δ 144.4 (C-6 or C-7), 143.7 (C-6 or C-7), 125.4 (C-5a), 123.9 (C-8a), 116.4 (C-5), 114.5 (C-8), 61.2 (C-1), 41.1 (C-3), 31.0 (C-1'), 24.8 (C-4), 19.0 (C-2'); *m/z* [ES+] 208 ([M+H]$^+$, 100%); *m/z* [HRMS ESI+] found [M+H]$^+$ 208.1336; $C_{12}H_{27}NO_2$ requires 208.1338; retention time (achiral analytical HPLC method 1) = 4.4 min.

(*S*)-1-((*R*)-*sec*-Butyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (94)[148]



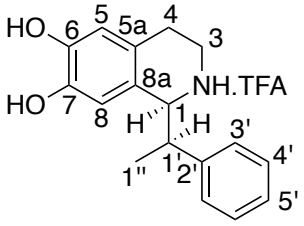
The product was isolated as a white solid (12 mg, 37%). $[\alpha]_D^{25}$ 0.017 (c = 0.32, MeOH); υ_{max}/ cm^{-1} (thin film): 3500-2500, 1670, 1615, 1527; ^{1}H-NMR (700 MHz; D_2O) δ 6.79 (1H, s, 8-H), 6.76 (1H, s, 5-H), 4.52 – 4.48 (0.95H, m, 1-H), 4.44 – 4.41 (0.05H, m, 1-H), 3.61 – 3.55 (1H, m, 3-*H*H), 3.28 – 3.20 (1H, m, 3-H*H*), 3.01 (1H, m, 4-*H*H), 2.88 (1H, m, 4-H*H*), 2.27 (1H, m, 1'-H), 1.53 (1H, m, 2'-H), 1.42 (1H, m, 2'-H), 1.09 (0.18H, d, *J* = 7.3 Hz, 1''-H), 1.02 (2.9H, t, *J* = 8.6 Hz, 3'-H), 0.89 (0.2H, t, *J* = 8.4 Hz, 3'-H), 0.80 (2.87H, d, *J* = 8.3 Hz, 3'-H); ^{13}C-NMR (176 MHz; D_2O) δ 144.3 (C-6 or C-7), 143.8 (C-6 or C-7), 125.7 (C-5a), 123.9 (C-8a), 116.3 (C-5), 114.1 (C-8), 59.8 (C-1), 41.7 (C-3), 37.9 (C-1'), 26.4 (C-4), 24.9 (C-2'), 12.6 (C-1''), 11.8 (C-3'); *m/z* [ES+] 222 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 222.1494; $C_{13}H_{20}NO_2$ requires 222.1494; retention time (achiral analytical HPLC method 1) = 5.0 min.

(*S*)-1-((*R*)-Pentan-2-yl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (95)[148]



The product was isolated as a white solid (3.0 mg, 9.0%). $[\alpha]_D^{25}$ -0.017 (c = 0.20, MeOH); υ_{max}/ cm^{-1} (thin film): 3500-2800, 2965, 2875, 1670, 1613, 1528; ^{1}H-NMR (700 MHz; D_2O) δ 6.79 – 6.75 (1H, m, 8-H), 6.73 (1H, s, 5-H), 4.45 (0.8H, d, *J* = 4.2 Hz, 1-H), 4.41 (0.2H, d, *J* = 3.7 Hz, 1-H), 3.55 - 3.51 (1H, m, 3-*H*H), 3.24 - 3.18 (1H, m, 3-H*H*), 3.00 - 2.93 (1H, m, 4-*H*H), 2.86 - 2.80 (1H, m, 4-H*H*), 2.28 – 2.36 (1H, m, 1'-H), 1.48 - 1.29 (3.6H, m, 2'-H and 3'-H), 1.19 - 1.09 (0.4H, m, 2'-H and 3'-H), 1.06 (0.7H, d, *J* = 6.8 Hz, 4'-H), 0.89 (2.3H, t, *J* = 7 Hz, 4'-H), 0.77 (3H, d, *J* = 7.2 Hz, 1''-H); ^{13}C-NMR (176 MHz; D_2O) δ 144.3 (C-6 or C-7), 143.8 (C-6 or C-7), 125.6 (C-8a), 123.9 (C-5a), 116.3 (C-5), 114.2 and 114.1 (C-8a), 61.2 and 60.0 (C-1), 41.6 and 41.5 (C-3), 35.7 and 35.4 (C-1'), 24.9 (C-4), 20.5 and 20.4 (C-3'), 13.8 and 13.7 (C-1''), 12.9 (C-4'); *m/z* [ES+] 236 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 236.1644; $C_{14}H_{22}NO_2$ requires 236.1644; retention time (achiral analytical HPLC method 1) = 5.6 min.[148]

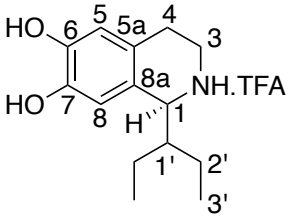
(*S*)-1-((*R*)-1-Phenylethyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (96)[148]



The product was isolated as a white solid (5.7 mg, 16%). $[\alpha]_D^{25}$ -0.011 (c = 0.12, MeOH); υ_{max}/ cm^{-1} (thin film): 3300-2500, 1672, 1610, 1538; ^{1}H-NMR (700 MHz; D_2O) δ 7.38 – 7.31 (2H, m, 4'-H), 7.30 – 7.27 (1H, m, 5'-H), 7.24 – 7.20 (2H, d, *J* = 7.0 Hz, 3'-H), 6.66 (1H, s, 8-H), 6.34 (1H, s, 5-H), 4.58 (0.9H, d, *J* = 6.3 Hz, 1-H), 4.55 (0.1 H, d, *J* = 9.1 Hz, 1'-H), 3.60 – 3.52 (1H, m, 3-*H*H), 3.45 - 3.39 (1H, m, 3-H*H*), 3.19 – 3.13 (1H, m, 4-

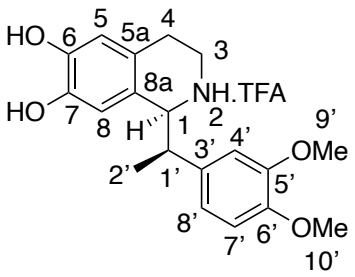
*H*H), 2.95 – 2.85 (1H, m, 4-H*H*), 2.85 – 2.79 (1H, m, 1'-H), 1.23 (3H, d, *J* = 7.0 Hz, 1''-H); ^{13}C-NMR (176 MHz; D_2O) δ 144.6 (C-2'), 143.2 (C-6 or C-7), 141.2 (C-6 or C-7), 129.6 (C-5a, C-8a or C-4'), 128.7 (C-5a, C-8a or C-4'), 128.3 (C-5a, C-8a or C-4'), 125.0 (C-3'), 123.0 (C-5'), 116.1 (C-5), 114.9 (C-8), 61.1 (C-1), 41.9 (C-3), 40.1 (C-1'), 24.4 (C-4), 14.1 (C-1''); *m/z* [ES+] 270 ([M+H]$^+$, 100%); *m/z* [HRMS ES$^+$] found [M+H]$^+$ 270.1488; $C_{17}H_{20}NO_2$ requires 270.1489; retention time (achiral analytical HPLC method 1) = 5.8 min.

(*S*)-1-(Pentan-3-yl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (98)[148]



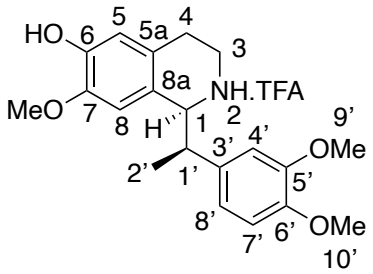
The product was isolated as a white solid (15 mg, 47%). $[\alpha]_D^{25}$ +0.39 (c = 0.35, MeOH); υ_{max}/ cm^{-1} (thin film): 3200-2800, 2500-1900, 1714, 1614; ^{1}H-NMR (700 MHz; D_2O) δ 6.74 (1H, s, 8-H), 6.69 (1H, s, 5-H), 4.53 (1H, d, *J* = 3.4 Hz, 1-H), 3.52 - 3.46 (1H, m, 3-*H*H), 3.22 - 3.15 (1H, m, 3-H*H*), 2.98 - 2.90 (1H, m, 4-*H*H), 2.84 - 2.77 (1H, m, 4-H*H*), 1.95 - 1.88 (1H, m), 1.64 - 1.55 (1H, m), 1.29 - 1.20 (1H, m), 1.20 - 1.13 (1H, m), 1.13 - 1.06 (1H, m), 0.94 (3H, t, *J* = 7.2 Hz, 3'-H), 0.77 (3H, t, *J* = 7.2 Hz, 3'-H); ^{13}C-NMR (176 MHz; D_2O) δ 144.2 (C-6 or C-7), 143.0 (C-6 or C-7), 125.8 (C-8a), 124.0 (C-5a), 116.3 (C-5), 114.1 (C-8), 57.8 (C-1), 44.8 (C-1'), 41.8 (C-3), 24.9 (C-4), 23.0 and 21.9 (C-2'), 12.3 and 11.6 (C-3'); *m/z* [ES+] 236 ([M+H]+, 100%); m/z [HRMS ES+] found [M+H]+ 236.1649; $C_{14}H_{22}NO_2$ requires 236.1651; retention time (achiral analytical HPLC method 1) = 5.1 min.

(*S*)-1-((*R*)-1-(3,4-Dimethoxyphenyl)ethyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (119)



A solution of dopamine hydrochloride (18.9 mg, 0.1 mmol) and sodium ascorbate (19.8 mg, 0.1 mmol) in HEPES buffer (50 mM, pH 7.5) was prepared. A solution of 2-(3,4-dimethoxyphenyl)propanal (38.8 mg, 0.2 mmol) in acetonitrile (1 mL) was prepared and the two solutions mixed. Δ29*Tf*NCS-M97V (2 mg at 10 mg mL^{-1} in 20 mM Tris, 50 mM NaCl, pH 7.5) was added and the reaction mixture stirred at 37 °C for 18 h. Once the reaction was completed (monitored by depletion of dopamine by HPLC method 1), workup method 5 was performed and the resulting residue purified by preparative-HPLC (method 9) to isolate the pure product (as the TFA salt) as white solid (7.8 mg, 18%) with a diastereomeric ratio of 95:5 (1*S*,1'*R*:1*S*:1'*S*)-**119**. $[\alpha]_D^{25}$ +15.2 (c = 0.71, MeOH); υ_{max}/ cm^{-1} (thin film): 2971, 2840, 1671, 1609, 1517; ^{1}H-NMR (700 MHz, CD_3OD) δ 6.94 (1H, d, *J* = 8.4 Hz, 7'-H), 6.82 (1H, dd, *J* = 8.4 Hz, 1.9 Hz, 8'-H), 6.78 – 6.75 (1H, m, 4'-H), 6.60 (1H, s, 5-H), 6.46 (1H, s, 8-H), 4.58 (1H, d, *J* = 6.1 Hz, 1-H), 3.82 (3H, s, 9'-OC*H*$_3$ or 10'-OC*H*$_3$), 3.77 (3H, s, 9'- OC*H*$_3$ or 10'- OC*H*$_3$), 3.53 – 3.48 (1H, m, 1-H), 3.42 – 3.37 (1H, m, 3-*H*H), 3.21 – 3.14 (1H, m, 3-H*H*), 2.92 – 2.82 (2H, m, 4-H), 1.31 (3H, d, *J* = 7.3 Hz, 2'-H); ^{13}C-NMR (176 MHz, CD_3OD) δ 150.7 (C-5'), 150.0 (C-6'), 146.8 (C-6 or C-7), 145.5 (C-6 or C-7), 134.5 (C-3'), 124.2 (C-5a), 122.5 (C-8a), 121.4 (C-8'), 116.0 (C-5), 114.6 (C-8), 113.1 (C-7' or C-4'), 113.0 (C-7' or C-4'), 61.9 (C-1), 56.3 (C-9'), 56.3 (C-10'), 43.1 (C-1'), 41.0 (C-3), 25.4 (C-4), 15.5 (C-2'), *m/z* [ES^+] 330 ([M+H]$^+$, 100%); m/z [HRMS ES+] found [M+H]+ 330.1700; $C_{19}H_{24}NO_4$ requires 330.1700; retention time (achiral analytical HPLC method 2): 7.9 min.

1-((1*S*)-1-((*R*)-1-(3,4-Dimethoxyphenyl)ethyl)-6-hydroxy-7-methoxy-1,2,3,4-tetrahydro-2λ^4-isoquinolin-2-yl)-2,2,2-trifluoroethan-1-one.TFA (127)

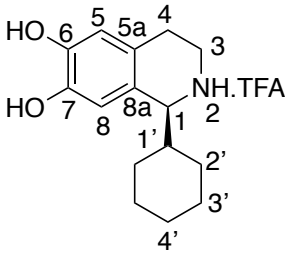


The same method was used as for the synthesis of **119,** except 4-methoxytyramine (0.1 mmol) was used as the phenethylamine substrate, rather than dopamine HCl (**5**). The product (TFA salt) was isolated by preparative HPLC (method 9) as a white solid (35.5 mg, 81%) with a diastereomeric ratio of 92:8 of (1*S*,1'*R*:1*S*:1'*S*)-**127.** $[\alpha]_D^{25}$ +1.5 (c = 2.2, MeOH); ν_{max}/ cm^{-1} (thin film): 2919, 1674, 1608, 1516; ^{1}H-NMR (600 MHz; CD_3OD) δ 6.96 (1H, d, *J* = 8.3 Hz, 7'-H), 6.82 (1H, dd, *J* = 8.3, 2.0 Hz, 8'-H), 6.79 (1H, d, *J* = 2.0 Hz, 4'-H), 6.63 (1H, s, 5-H), 6.20 (1H, s, 8-H), 4.85 (1H, d, *J* = 7.7 Hz, 1-H), 3.82 (3H, s, 9'- OC*H*$_3$ or 10'- OC*H*$_3$), 3.77 (3H, s, 9'- OC*H*$_3$ or 10'- OC*H*$_3$), 3.55 (3H, s, 7-OC*H*$_3$), 3.54 – 3.42 (2H, m, 1'-H and 3-*H*H), 3.30 – 3.25 (1H, m, 3-H*H*), 3.02 – 2.91 (2H, m, 4-H), 1.38 (3H, d, *J* = 7.1 Hz, 2'-H); ^{13}C-NMR (151 MHz; CD_3OD) δ 150.7 (C-5' or C-6'), 150.0 (C-5' or C-6'), 147.8 (C-6 or C-7),147.7 (C-6 or C-7), 135.1 (C-3'), 125.4 (C-5a), 122.8 (C-8a), 121.6 (C-8'), 116.0 (C-5), 113.5 (C-4'), 113.2 (C-7'), 111.7 (C-8), 61.9 (C-1), 56.5 (O*C*H$_3$), 56.5 (O*C*H$_3$), 56.2 (O*C*H$_3$), 43.5 (C-1'), 40.1 (C-3), 25.4 (C-4), 16.3 (C-2'); *m/z* [ES$^+$] 344 ([M+H]$^+$, 100%); m/z [HRMS ES+] found [M+H]+ 344.1862; $C_{20}H_{26}NO_4$ requires 344.1856.

8.7.3 Products from reactions with benzaldehydes

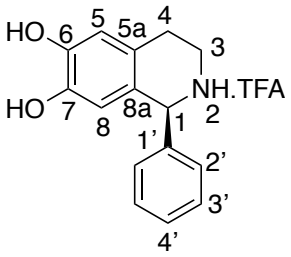
Preparative scale reactions were performed on a 10 mL scale (10 mM amine, 20 mM aldehyde) using wild-type *Tf*NCS (0.2 mg mL^{-1} final concentration of purified enzyme) for 3 h. Workup method 5 was used followed by preparative HPLC purification.

(1*S*)-1-Cyclohexyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (156)[149]



Purified by preparative HPLC (method 9) and isolated as a white solid (18 mg, 52%). mp > 250 °C (H_2O); $[\alpha]_D^{25}$ 8.2 (*c* 0.38, MeOH); ν_{max}/ cm^{-1} (thin film): 3354, 2933, 2835, 1622, 1530; ^{1}H-NMR (700 MHz; D_2O) δ 6.73 (1H, s, 5-H or 8-H), 6.71 (1H, s, 5-H or 8-H), 4.26 (1H, d, *J* = 5.0 Hz, 1-H), 3.54 – 3.48 (1H, dt, *J* = 12.3, 5.3 Hz, 3-*H*H), 3.25 – 3.19 (1H, m, 3-H*H*), 2.97 – 2.90 (1H, m, 4-*H*H), 2.84 (1H, dt, *J* = 16.8, 5.3 Hz, 4-H*H*), 2.04 – 1.98 (1H, m, 1'-H), 1.78 – 1.73 (1H, m), 1.71 – 1.66 (2H, m), 1.65 – 1.61 (1H, m), 1.48 – 1.43 (1H, m), 1.29 – 1.05 (4H, m), 0.97 (1H, qd, *J* = 12.6, 3.6 Hz); ^{13}C-NMR (176 MHz; CD_3OD) δ 147.0 (C-6 or C-7), 145.9 (C-6 or C-7), 138.8 (C-1' or C-6'), 136.3 (C-1' or C-6'), 132.2 (C-4'), 130.7 (C-5'), 130.3 (C-2'), 128.0 (C-3'), 124.2 (C-5a or C-8a), 123.8 (C-5a or C-8a), 115.9 (C-5), 114.9 (C-8), 57.2 (C-1), 42.0 (C-3), 25.6 (C-4), 19.3 (C-7'); *m/z* [ES+] 248 ($[M+H]^+$, 100%); *m/z* [HRMS ES+] found $[M+H]^+$ 248.1645; $C_{15}H_{21}NO_2$ requires 248.1645; retention time (achiral analytical HPLC method 1): 5.4 min.

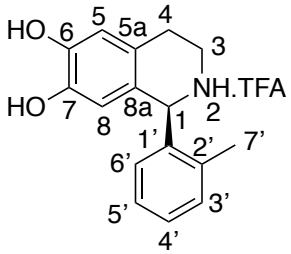
(1*S*)-1-Phenyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (157)[149]



Purified by preparative HPLC (method 9) and isolated as a white solid (17 mg, 51%). mp > 250 °C (H_2O); $[\alpha]_D^{25}$ 6.5 (*c* 0.37, MeOH); ν_{max}/ cm^{-1} (thin film): 3033, 2852, 1668, 1612, 1527; ^{1}H-NMR (700 MHz; D_2O) δ 7.49 – 7.43 (3H, m, 3'-H and 4'-H), 7.34 – 7.29

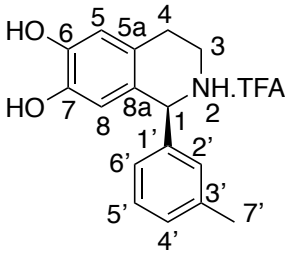
(2H, m, 2'-H), 6.81 (1H, s, 8-H), 6.34 (1H, s, 5-H), 5.63 (1H, s, 1-H), 3.50 – 3.39 (2H, m, 3-H), 3.10 – 3.04 (2H, m, 4-H); ^{13}C-NMR (176 MHz, D_2O) δ 147.1 (C-6 or C-7), 145.9 (C-6 or C-7), 138.0 (C-1'), 130.8 (C-2'), 130.8 (C-3'), 130.2 (C-4'), 124.2 (C-5a), 123.1 (C-8a), 115.8 (C-8), 115.2 (C-5), 60.9 (C-1), 41.4 (C-3), 25.5 (C-4); *m/z* [ES+] 242 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 242.1176; $C_{15}H_{15}NO_2$ requires 242.1176; retention time (achiral analytical HPLC method 1): 4.9 min.

(1*S*)-1-(*o*-Tolyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (158)[149]



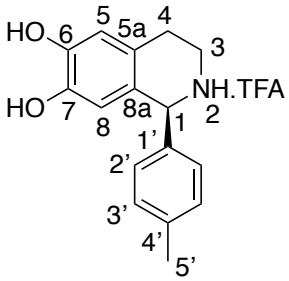
Purified by preparative HPLC (method 9) and isolated as a white solid (8.7 mg, 25%); $[\alpha]_D^{25}$ 8.8 (*c* 0.43, MeOH); ν_{max}/ cm^{-1} (thin film): 3024, 2818, 1663, 1609, 1526; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.38 – 7.30 (2H, m, 4'-H and 5'-H), 7.29 – 7.24 (1H, m, 3'-H), 7.10 (1H, d, *J* = 7.8 Hz, 2'-H), 6.68 (1H, s, 5-H), 6.08 (1H, s, 8-H), 5.80 (1H, s, 1'-H), 3.56 – 3.47 (2H, m, 3-H), 3.21 – 3.15 (1H, m, 4-*H*H), 3.04 – 2.97 (1H, dt, *J* = 17.2, 5.3 Hz, 4-H*H*), 2.51 (3H, s, 7'-H); ^{13}C-NMR (176 MHz; CD_3OD) δ 147.0 (C-6 or C-7), 145.9 (C-6 or C-7), 138.8 (C-1' or C-6'), 136.3 (C-1' or C-6'), 132.2 (C-4'), 130.7 (C-5'), 130.3 (C-2'), 128.0 (C-3'), 124.2 (C-5a or C-8a), 123.8 (C-5a or C-8a), 115.9 (C-5), 114.9 (C-8), 57.2 (C-1), 42.0 (C-3), 25.6 (C-4), 19.3 (C-7'); *m/z* [ES+] 256 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 256.1331; $C_{16}H_{18}NO_2$ requires 256.1332; retention time (achiral analytical HPLC method 1): 5.2 min.

(1*S*)-1-(*m*-Tolyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (159)[149]



Purified by preparative HPLC (method 9) and isolated as a white solid (4.3 mg, 12%). $[\alpha]_D^{25}$ 6.1 (*c* 0.18, MeOH); ν_{max}/ cm^{-1} (thin film): 3436, 2977, 2809, 1665, 1607, 1527; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.36 (1H, t, *J* = 7.7 Hz, 5'-H), 7.29 (1H, d, *J* = 7.7 Hz, 4'-H), 7.18 (1H, s, 2'-H), 7.15 (1H, d, *J* = 7.7 Hz, 6'-H), 6.68 (1H, s, 5-H), 6.19 (1H, s, 8-H), 5.51 (1H, s, 1-H), 3.50 – 3.45 (1H, m, 3-*H*H), 3.43 – 3.38 (1H, m, 3-H*H*), 3.17 – 3.11 (1H, m, 4-*H*H), 3.00 (1H, dt, *J* = 17.2, 5.3 Hz, 4-H*H*), 2.37 (3H, s, 7'-H); ^{13}C-NMR (176 MHz; CD_3OD) δ 147.1 (C-6 or C-7), 145.8 (C-6 or C-7), 140.3 (C-1' or C-3'), 138.0 (C-1' or C-3'), 131.5 (C-4'), 131.2 (C-2'), 130.1 (C-3'), 127.8 (C-6'), 124.2 (C-5a or C-8a), 123.2 (C-5a or C-8a), 115.8 (C-5), 115.2 (C-8), 60.9 (C-1), 41.4 (C-3), 25.6 (C-4), 21.2 (C-7'); *m/z* [ES+] 256 ($[M+H]^+$, 100%); *m/z* [HRMS ES+] found $[M+H]^+$ 256.1332; $C_{16}H_{18}NO_2$ requires 256.1332; retention time (achiral analytical HPLC method 1): 5.3 min.

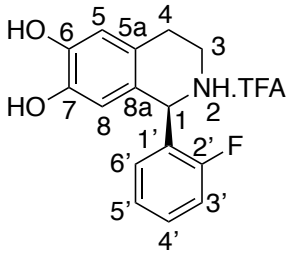
(1*S*)-1-(*p*-Tolyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (160)[149]



Purified by preparative HPLC (method 9) and isolated as a white solid (3.8 mg, 11%). $[\alpha]_D^{25}$ 1.3 (*c* 0.091, MeOH)**;** ^{1}H-NMR (700 MHz; CD_3OD) δ 7.29 (2H, d, *J* = 8.0 Hz, 3'-H), 7.23 (2H, d, *J* = 8.0 Hz, 2'-H), 6.67 (1H, s, 5-H), 6.19 (1H, s, 8-H), 5.51 (1H, s, 1'-H), 3.48 – 3.43 (1H, m, 3-*H*H), 3.42 – 3.37 (1H, m, 3-H*H*), 3.16 – 3.09 (1H, m, 4-*H*H),

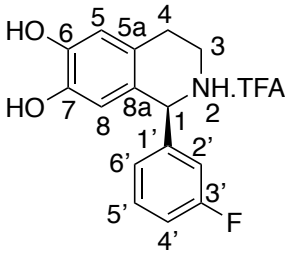
3.99 (1H, dt, *J* = 17.2, 5.5 Hz, 4-*H*H), 2.37 (3H, s, 5'-H); ^{13}C-NMR (176 MHz; CD_3OD) δ 147.0 (C-6 or C-7), 145.8 (C-6 or C-7), 141.1 (C-4'), 135.1 C-1'), 130.8 (C-3'), 130.6 (C-2'), 124.2 (C-5a), 123.4 (C-8a), 115.8 (C-5), 115.2 (C-8), 60.7 (C-1), 41.3 (C-3), 25.6 (C-4), 21.1 (C-5'); *m/z* [ES+] 256 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 256.1331; $C_{16}H_{18}NO_2$ requires 256.1332; retention time (achiral analytical HPLC method 1): 5.3 min.

(1*S*)-1-(2-Fluorophenyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (161)[149]



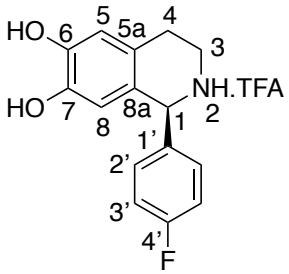
Purified by preparative HPLC (method 9) and isolated as a white solid (11 mg, 31%). $[\alpha]_D^{25}$ -5.2 (*c* 0.23, MeOH); ν_{max}/ cm^{-1} (thin film): 3116, 2317, 1670, 1611, 1529; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.52 (1H, d, *J* = 5.3 Hz, 3'-H), 7.31 – 7.24 (2H, m, 3'-H and 4'-H), 7.19 (1H, d, *J* = 7.3 Hz, 6'-H), 6.70 (1H, s, 5-H), 6.20 (1H, s, 8-H), 5.87 (1H, s, 1-H), 3.46 (2H, app.t, *J* = 6.1 Hz, 3-H), 3.16 – 3.09 (1H, m, 4-*H*H), 3.02 (1H, dt, *J* = 17.2, 5.5 Hz, 4-H*H*); ^{13}C-NMR (176 MHz; CD_3OD) δ 162.4 (d, $^1J_{CF}$ = 248 Hz, C-2'), 147.2 (C-6 or C-7), 146.0 (C-6 or C-7), 133.2 (d, J_{CF} = 8.4 Hz, C-6'), 132.5 (d, J_{CF} = 2.1 Hz, C-1'), 126.1 (d, J_{CF} = 3.5 Hz, C-4'), 125.0 (d, $^3J_{CF}$ = 12 Hz, C-5'), 124.4 (C-5a), 121.9 (C-8a), 117.1 (d, $^2J_{CF}$ = 21 Hz, C-3'), 116.0 (C-5), 114.6 (C-8), 54.0 (C-1), 41.4 (C-3), 25.5 (C-4); *m/z* [ES+] 260 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 260.1081; $C_{15}H_{15}FNO_2$ requires 260.1081; retention time (achiral analytical HPLC method 1): 4.7 min.

(1*S*)-1-(3-Fluorophenyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (162)[149]



Purified by preparative HPLC (method 9) and isolated as a white solid (11.0 mg, 31%); $[\alpha]_D^{25}$ -2.9 (*c* 0.41, MeOH); ν_{max}/ cm^{-1} (thin film): 3072, 2558, 2324, 1667, 1612, 1592, 1519; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.51 (1H, dd, *J* = 5.8 Hz, 5'-H), 7.25 – 7.18 (2H, m, 4'-H and 6'-H), 7.13 (1H, d, *J* = 9.2 Hz, 2'-H), 6.69 (1H, s, 5-H), 6.20 (1H, s, 8-H), 5.62 (1H, s, 1-H), 3.52 – 3.37 (2H, m, 3-H), 3.18 – 3.09 (1H, m, 4-*H*H), 3.04 – 2.96 (1H, m, 4-H*H*); ^{13}C-NMR (176 MHz; CD_3OD) δ 165.0 (d, $^1J_{CF}$ = 245 Hz, C-3'), 147.3 (C-6 or C-7), 145.9 (C-6 or C-7), 140.5 (d, J_{CF} = 7.0 Hz, C-1'), 132.1 (d, J_{CF} = 8.3 Hz, C-5'), 126.8 (d, J_{CF} = 2.8 Hz, C-6'), 124.3 (C-5a), 122.6 (C-8a), 117.7 (dd, $^2J_{CF}$ = 21.9, 4.8 Hz, C-2'), 115.9 (C-5), 115.1 (C-8), 60.1 (C-1), 41.4 (C-3), 25.5 (C-4); *m/z* [ES+] 260 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 260.1081; $C_{15}H_{15}FNO_2$ requires 260.1081; retention time (achiral analytical HPLC method 1): 5.0 min.

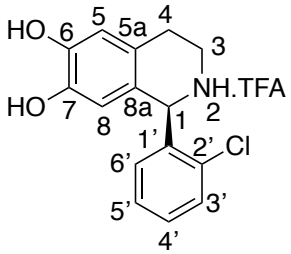
(1*S*)-1-(4-Fluorophenyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (163)[149]



Purified by preparative HPLC (method 9) and isolated as a white solid (10 mg, 29%). $[\alpha]_D^{25}$ -5.4 (*c* 0.46, MeOH); ν_{max}/ cm^{-1} (thin film): 3112, 2819, 1668, 1606, 1560, 1532, 1508; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.43 – 7.37 (2H, m, 2'-H), 7.25 – 7.18 (2H, m, 3'-H), 6.69 (1H, s, 5-H), 6.19 (1H, s, 8-H), 5.61 (1H, s, 1-H), 3.49 – 3.44 (1H, m, 3-*H*H), 3.44 – 3.28 (1H, m, 3-H*H*), 3.16 – 3.08 (1H, m, 4-*H*H), 3.02 (1H, dt, *J* = 17.2, 5.7 Hz, 4-

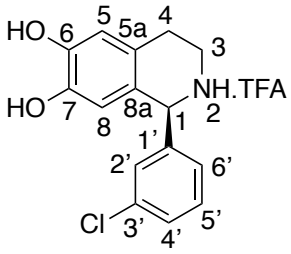
H*H*); ^{13}C-NMR (176 MHz; CD_3OD) δ 164.7 (d, $^1J_{CF}$ = 251 Hz, C-4'), 147.2 (C-6 or C-7), 145.9 (C-6 or C-7), 134.1 (d, J_{CF} = 3.1 Hz, C-1'), 133.1 (d, J_{CF} = 8.5 Hz, C-2'), 124.3 (C-5a), 122.9 (C-8a), 117.0 (d, $^2J_{CF}$ = 23 Hz, C-3'), 115.7 (C-5), 115.1 (C-8), 60.0 (C-1), 41.2 (C-3), 25.5 (C-4); *m/z* [ES+] 260 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 260.1081; $C_{15}H_{15}FNO_2$ requires 260.1081; retention time (achiral analytical HPLC method 1): 5.1 min.

(1*S*)-1-(2-Chlorophenyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (164)[149]



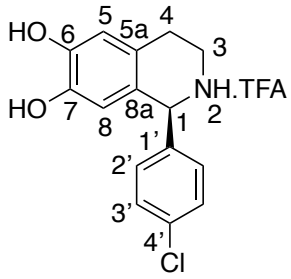
Purified by preparative HPLC (method 9) and isolated as a white solid (13 mg, 35%). $[\alpha]_D^{25}$ 2.0 (*c* 0.12, MeOH); ν_{max}/ cm^{-1} (thin film): 2995, 2816, 2580, 1663, 1609, 1528; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.61 (1H, dd, *J* = 8.1, 1.1 Hz, 3'-H), 7.48 (1H, td, *J* = 7.7, 1.6 Hz), 7.39 (1H, td, *J* = 7.7, 1.1 Hz, 5'-H or 4'-H), 7.19 (1H, dd, *J* = 7.7, 1.6 Hz), 6.70 (1H, s, 5-H), 6.14 (1H, s, 8-H), 6.04 (1H, s, 1-H), 3.50 – 3.43 (2H, m, 3-H), 3.11 (1H, dt, *J* = 17.3, 6.4 Hz, 4-*H*H), 3.05 – 3.00 (1H, *J* = 17.3, 6.2 Hz, 4-H*H*); ^{13}C-NMR (176 MHz, CD_3OD) δ 147.3 (C-6 or C-7), 146.2 (C-6 or C-7), 135.8 (C-1' or C-2'), 135.6 (C-1' or C-2'), 132.6 (C-6' or C-5'), 132.5 (C-6' or C-5'), 131.3 (C-3'), 129.0 (C-4'), 124.5 (C-5a), 122.1 (C-8a), 115.9 (C-5), 114.7 (C-8), 56.9 (C-1), 41.2 (C-3), 25.5 (C-4); *m/z* [ES+] 276 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 276.0786; $C_{15}H_{15}ClNO_2$ requires 276.0786; retention time (achiral analytical HPLC method 1): 5.1 min.

(1*S*)-1-(3-Chlorophenyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (165)[149]



Purified by preparative HPLC (method 9) and isolated as a white solid (5.2 mg, 14%). $[\alpha]_D^{25}$ 1.5 (*c* 0.015, MeOH); ν_{max}/ cm^{-1} (thin film): 3028, 2820, 2343, 2228, 1664, 1604, 1528; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.52 – 7.45 (2H, m, 4'-H and 5'-H), 7.41 (1H, s, 2'-H), 7.30 (1H, d, *J* = 7.7 Hz, 6'-H), 6.69 (1H, s, 8-H), 6.19 (1H, s, 5-H), 5.60 (1H, s, 1-H), 3.51 – 3.38 (2H, m, 3-H), 3.17 – 3.10 (1H, m, 4-*H*H), 3.01 (1H, dt, *J* = 17.2, 5.5 Hz, 4-H*H*); ^{13}C-NMR (176 MHz; CD_3OD) δ 147.3 (C-6 or C-7), 146.0 (C-6 or C-7), 140.2 (C-1'), 136.1 (C-3'), 131.8 (C-4'), 130.9 (C-2' or C-5'), 130.8 (C-2' or C-5'), 129.3 (C-6'), 124.3 (C-5a), 122.4 (C-8a), 115.9 (C-5), 115.0 (C-8), 60.1 (C-1), 41.4 (C-3), 25.5 (C-4); *m/z* [ES+] 276 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 276.0786; $C_{15}H_{15}ClNO_2$ requires 276.0786; retention time (achiral analytical HPLC method 1): 5.4 min.

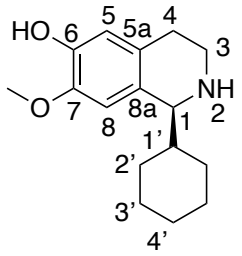
(*S*)-1-(4-Chlorophenyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (166)[149]



Purified by preparative HPLC (method 9) and isolated as a white solid (6.9 mg, 19%). $[\alpha]_D^{25}$ 1.1 (*c* 0.29, MeOH); ν_{max}/ cm^{-1} (thin film): 3166, 3102, 2772, 2634, 2556, 1800, 1670, 1610, 1589, 1535; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.49 (2H, d, *J* = 8.5 Hz, 3'-H), 7.35 (2H, d, *J* = 8.5 Hz, 2'-H), 6.68 (1H, s, 5-H), 6.18 (1H, s, 8-H), 5.59 (1H, s, 1-H), 3.49 – 3.37 (2H, m, 3-H), 2.16 – 3.01 (1H, m, 4-*H*H), 3.00 (1H, dt, *J* = 17.2, 5.5 Hz, 4-H*H*); ^{13}C-NMR (176 MHz; CD_3OD) δ 147.2 (C-6 or C-7), 146.0 (C-6 or C-7), 136.8 (C-

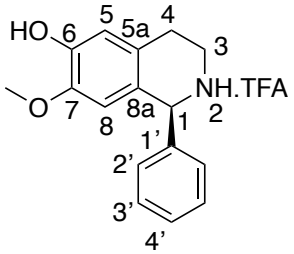
1'), 136.8 (C-4'), 132.5 (C-2'), 130.3 (C-3'), 124.3 (C-5a), 122.7 (C-8a), 115.9 (C-5), 115.1 (C-8), 60.0 (C-1), 41.2 (C-3), 25.5 (C-4); *m/z* [ES+] 276 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 276.0786; $C_{15}H_{15}ClNO_2$ requires 276.0786; retention time (achiral analytical HPLC method 1): 4.8 min.

(1*S*)-1-Cyclohexyl-7-methoxy-1,2,3,4-tetrahydroisoquinolin-6-ol.TFA (173)[149]



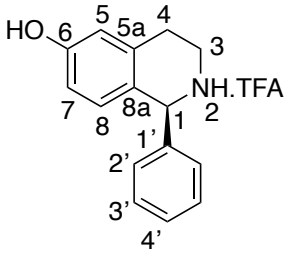
Synthesised from NCS mediated reaction between 4-methoxytyramine hydrochloride and cyclohexanecarboxaldehyde. Purified by preparative HPLC (method 9) and isolated as a white solid (8.4 mg, 32%). $[\alpha]_D^{26}$ -1.1 (*c* 0.29, MeOH); ν_{max}/ cm^{-1} (thin film): 2971, 2853, 2165, 2033, 1677, 1517; ^{1}H-NMR (700 MHz; CD_3OD) δ 6.76 (1H, s, 8-H), 6.65 (1H, s, 5-H), 4.31 (1H, d, *J* = 5.6 Hz, C-1), 3.83 (3H, s, OC*H*$_3$), 3.53 (1H, dt, *J* = 12.3, 5.6 Hz, 3-*H*H), 3.28 – 3.22 (1H, m, 3-H*H*), 3.01 – 2.95 (1H, m, 4-*H*H), 2.88 (1H, dt, *J* = 16.9, 5.3 Hz, 4-H*H*), 2.12 – 2.06 (1H, m, 1'-H), 1.89 – 1.84 (1H, m), 1.83 – 1.71 (3H, m), 1.59 – 1.54 (1H, m), 1.43 – 1.35 (1H, qt, *J* = 12.8, 3.6 Hz), 1.35 – 1.26 (1H, m), 1.26 – 1.18 (1H, m), 1.10 (1H, qd, *J* = 12.6, 3.6 Hz); ^{13}C-NMR (176 MHz; CD_3OD) δ 148.3 (C-6 or C-7), 147.7 (C-6 or C-7), 125.9 (C-8a), 122.6 (C-5a), 116.2 (C-5), 111.0 (C-8), 61.5 (C-1), 56.5 (O*C*H$_3$), 41.9 (C-1'), 41.2 (C-3), 30.9 (C-2', C-3' or C-4'), 27.8 (C-2', C-3' or C-4'), 27.2 (C-2', C-3' or C-4'), 27.0 (C-2', C-3' or C-4'), 26.9 (C-2', C-3' or C-4'), 25.5 (C-4); *m/z* [ES+] 262 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 162.1799; $C_{16}H_{23}NO_2$ requires 262.1802.

(1*S*)-7-Methoxy-1-phenyl-1,2,3,4-tetrahydroisoquinolin-6-ol.TFA (174)[149]



Synthesised from NCS mediated reaction between 4-methoxytyramine hydrochloride and benzaldehyde. Purified by preparative HPLC (method 9) and isolated as white solid (12 mg, 24%). $[\alpha]_D^{26}$ -2.0 (*c* 0.35, MeOH)**;** ν_{max}/ cm^{-1} (thin film): 2925, 2853, 1631, 1515; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.49 – 7.47 (3H, m, 2'-H and 4'-H), 7.38 – 7.35 (2H, m, 3'-H), 6.73 (1H, s, 5-H), 6.31 (1H, s, 8-H), 5.67 (1H, s, 1-H), 3.61 (3H, s, OC*H*$_3$), 3.49 – 3.39 (2H, m, 3-H), 3.17 – 3.12 (1H, m, 4-*H*H), 3.07 – 3.01 (1H, m, 4-H*H*); ^{13}C-NMR (176 MHz; CD_3OD) δ 148.5 (C-6 or C-7), 148.2 (C-6 or C-7), 137.9 (C-1'), 130.9 (C-2' or C-3'), 130.9 (C-2' or C-3'), 130.3 (C-4'), 125.8 (C-5a), 122.8 (C-8a), 115.9 (C-5), 111.7 (C-8), 60.6 (C-1), 56.2 (O*C*H$_3$), 40.7 (C-3), 25.5 (C-4); *m/z* [ES+] 256 ([M+H]$^+$, 100%)); *m/z* [HRMS ES+] found [M+H]$^+$ 256.1333; $C_{16}H_{18}NO_2$ requires 256.1332.

(1*S*)-1-Phenyl-1,2,3,4-tetrahydroisoquinolin-6-ol.TFA (186)[149]



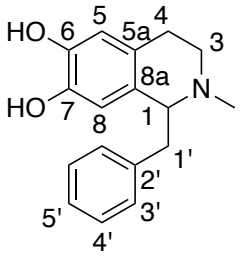
Purified by preparative HPLC (method 9) and isolated as a white solid (11 mg, 20%). $[\alpha]_D^{26}$ +5.5 (*c* 0.31, MeOH); ν_{max}/ cm^{-1} (thin film): 3068, 2323, 1666, 1613, 1591, 1529; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.49 – 7.46 (3H, m, 3'-H and 4'-H), 7.37 – 7.33 (2H, m, 2'-H), 6.72 (1H, s, 5-H), 6.65 – 6.63 (2H, m, 7-H and 8-H), 5.64 (1H, s, 1-H), 3.52 – 3.42 (2H, m, 3-H), 3.25 – 3.19 (1H, m, 4-*H*H), 3.10 (1H, dd, *J* = 17.6, 8.7 Hz, 4-H*H*); ^{13}C-NMR (176 MHz; CD_3OD) δ 158.7 (C-6), 137.9 (C-8a), 134.5 (C-5a), 130.9, 130.7, 130.4, 130.3, 123.1 (C-1'), 115.8 (C-5 or C-7), 115.7 (C-5 or C-7), 60.9 (C-1), 41.0 (C-

3), 26.2 (C-4); *m/z* [ES+] 226 ([M+H]$^+$, 100%)); *m/z* [HRMS ES+] found [M+H]$^+$ 226.1226; $C_{15}H_{16}NO$ requires 226.1226; retention time (achiral analytical HPLC method 1): 5.8 min.

8.7.4 Products from NCS mediated reactions with *N*-methyl dopamine

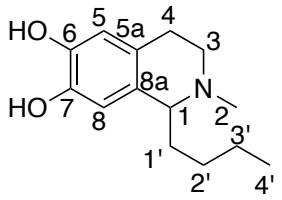
Preparative scale reactions were performed on a 10 mL scale (10 mM amine, 20 mM aldehyde) using wild-type *Tf*NCS (0.2 mg mL^{-1} final concentration of purified enzyme) for 18 h. Workup method 5 was used followed by preparative HPLC purification (method 11). To remove any residual TFA salts, workup method 3 was performed on the isolated products.

1-Benzyl-2-methyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol (203)



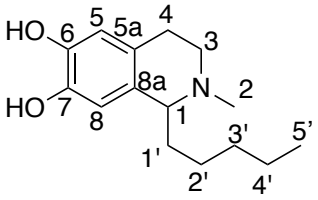
The product was isolated as a white solid (4.1 mg, 15%). ν_{max}/ cm^{-1} (thin film): 3037, 2756, 1673, 1612, 1530; ^{1}H-NMR (700 MHz; D_2O) δ 7.39-7.33 (3H, m, 5'-H and, 3'-H or 4'-H), 7.17 (2H, m, 3'-H or 4'-H), 6.76 (1H, s, 8-H), 6.17 (1H, s, 5-H), 4.58 (1H, t, *J* = 7.7 Hz, 1-H), 3.70 (1H, m, 1'-H), 3.38 - 3.28 (2H, m, 4-H), 3.20 (1H, m, 1'-H), 3.05 (1H, m, 3-*H*H), 2.97 (1H, m, 3-H*H*), 2.88 (3H, s, 2-H); ^{13}C-NMR (176 MHz; D_2O) δ 145.0 (C-6 or C-7), 143.2 (C-6 or C-7), 135.3 (C-2'), 130.5, 130.3, 129.6, 129.4, 128.4, 122.5, 121.5, 116.3, 116.0 (C-8), 115.9 (C-5), 65.3 (C-1), 47.2 (C-3), 45.5 (C-2), 40.4 (C-1'), 21.0 (C-4); *m/z* [ES+] 270 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 270.1495, $C_{17}H_{20}NO_2$ requires 270.1494; retention time (achiral analytical HPLC method 1) = 5.3 min.

1-Butyl-2-methyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol (204)



The product was isolated as a white solid (3.3 mg, 14%). ν_{max}/ cm^{-1} (thin film): 3184, 2957, 2869, 2724, 1619, 1526; ^{1}H-NMR (700 MHz; D_2O) δ 6.62 – 6.52 (2H, m, 5-H and 8-H), 4.15 – 4.06 (1H, m, 1-H), 3.59 - 3.49 (0.8H, m, 3-*H*H or 4-*H*H), 3.37 - 3.33 (0.5H, m, 3-*H*H or 4-*H*H), 3.19 – 3.13 (0.8H, m, 3-H*H* or 4-H*H*), 2.91 – 2.78 (2H, m, 3-H*H* or 4-H*H*), 2.73 (3H, m, 2-H), 1.80 – 1.61 (2H, m, 1'-H), 1.27 – 1.13 (4H, m, 2'-H and 3'-H), 0.74 – 0.69 (3H, m, 4'-H); ^{13}C-NMR (176 MHz, D_2O) δ 144.8 (C-6), 143.7 (C-7), 123.0 (C-8a), 122.6 (C-5a), 116.3 (C-5), 115.4 (C-8), 64.7 (C-1), 46.4 (C-3), 40.5 (C-2), 34.6 (C-1'), 27.3 (C-4), 22.4 (C-3), 13.5 (C-4'); *m/z* [ES+] 236 ([M+H]$^+$, 100%); *m/z* [HRMS ES$^+$] found [M+H]$^+$ 236.1646; $C_{14}H_{22}NO_2$ requires 236.1645; retention time (achiral analytical HPLC method 1) = 5.1 min.

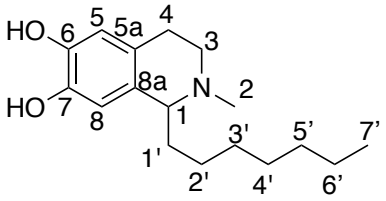
2-Methyl-1-pentyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol (205)



The product was isolated as a white solid (2.1 mg, 8.4%). ν_{max}/ cm^{-1} (thin film): 3141, 2954, 2930, 2859, 2726, 1616, 1526; ^{1}H-NMR (700 MHz; D_2O) δ 6.71 (1H, s, 8-H), 6.67 – 6.62 (1H, m, 5-H), 4.25 – 4.16 (1H, m, 1-H), 3.66 – 3.60 (0.8H, m, 3-*H*H or 4-*H*H), 3.48 - 3.43 (0.5H, m, 3-*H*H or 4-*H*H), 3.29 – 3.24 (0.8H, m, 3-*H*H or 4-*H*H), 3.03 – 2.89 (2H, m, 3-H*H* or 4-H*H*), 2.84 (3H, s, 2-H), 1.91 – 1.68 (2H, m, 1'-H), 1.39 – 1.18 (6H, m, 2'-H, 3'-H and 4'-H), 0.82 – 0.76 (3H, m, 5'-H); ^{13}C-NMR (176 MHz; D_2O) δ 144.8 (C-6), 143.6 (C-7), 123.0 (C-8a), 122.6 (C-5a), 116.3 (C-5), 115.4 (C-8), 64.7 (C-1), 46.4 (C-3), 40.5 (C-2), 34.8 (C-1'), 31.2 (C-4), 22.2 (C-3), 21.7 (C-4'), 13.7 (C-5'); *m/z*

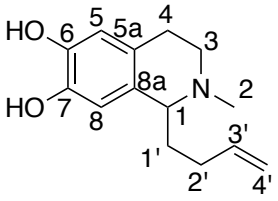
[ES+] 250 ([M+H]$^+$, 100%); *m/z* [HRMS ES$^+$] found [M+H]$^+$ 250.1804; $C_{15}H_{24}NO_2$ requires 250.1802; retention time (achiral analytical HPLC method 1) = 5.6 min.

1-Heptyl-2-methyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol (206)



The product was isolated as a white solid (2.8 mg, 10%). ν_{max}/ cm^{-1} (thin film): 3124, 2953, 2925, 2854, 2715, 2641, 1619, 1525; ^{1}H-NMR (700 MHz; D_2O) δ 6.72 (1H, s, 8-H), 6.67 – 6.64 (1H, m, 5-H), 4.25 – 4.18 (1H, m, 1-H), 3.67 – 3.61 (0.8H, m, 3-*H*H or 4-*H*H), 3.47 - 3.42 (0.5H, m, 3-*H*H or 4-*H*H), 3.28 – 3.23 (0.8H, m, 3-*H*H or 4-*H*H), 3.02 – 2.89 (2H, m, 3-H*H* or 4-H*H*), 2.85 – 2.81 (3H, s, 2-H), 1.92 – 1.71 (2H, m, 1'-H), 1.37 – 1.11 (10H, m, 2'-H, 3'-H, 4'-H, 5'-H and 6'-H), 0.76 (3H, t, *J* = 7.0 Hz, 7'-H); ^{13}C-NMR (176 MHz, D_2O) δ 144.8 (C-6), 143.6 (C-7), 123.0 (C-8a), 122.8 (C-5a), 116.3 (C-5), 115.5 (C-8), 64.8 (C-1), 46.4 (C-3), 40.6 (C-2), 34.7 (C-1'), 31.4 (C-4), 28.8 (C-2'), 28.5 (C-3'), 24.9 (C-4'), 22.5 (C-5'), 21.8 (C-6'), 13.9 (C-7'); *m/z* [ES+] 278 ([M+H]$^+$, 100%); *m/z* [HRMS ES$^+$] found [M+H]$^+$ 278.2119; $C_{17}H_{28}NO_2$ requires 278.2115; retention time (achiral analytical HPLC method 1) = 6.5 min

1-(But-3-en-1-yl)-2-methyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol (207)

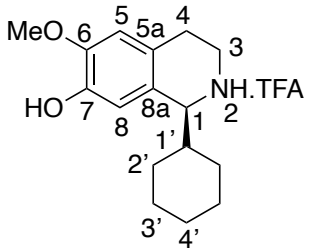


The product was isolated as a white solid (5.3 mg, 23%). ν_{max}/ cm^{-1} (thin film): 3301, 3105, 2954, 2693, 2637, 1641, 1610, 1525; ^{1}H-NMR (700 MHz; D_2O) δ 6.73 (1H, s, 8-H), 6.68 – 6.63 (1H, m, 5-H), 5.86 – 5.77 (1H, m, H-3'), 5.12 – 5.03 (2H, m, 5.12 – 5.03), 4.27 – 4.18 (1H, m, 1-H), 3.69 – 3.63 (0.8H, m, 3-*H*H or 4-*H*H), 3.50 - 3.46 (0.4H, m, 3-

*H*H or 4-*H*H), 3.31 – 3.19 (1H, m, 3-*H*H or 4-*H*H), 3.04 – 2.87 (2H, m, 3-H*H* or 4-H*H*), 2.14 – 2.09 (2H, m, 2'-H), 2.05 – 1.78 (2H, m, 1'-H); ^{13}C-NMR (176 MHz; D_2O) δ 144.9 (C-6), 143.6 (C-7), 137.4 (C-3'), 122.6 (C-8a), 122.5 (C-5a), 116.6 (C-4' or C-5), 116.4 (C-4' or C-5) 115.5 (C-8), 63.9 (C-1), 46.2 (C-3), 40.4 (C-2), 34.3 (C-1'), 29.4 (C-4), 21.4 (C-2'); *m/z* [ES+] 234 ([M+H]$^+$, 100%); *m/z* [HRMS ES$^+$] found [M+H]$^+$ 234.1480; $C_{14}H_{20}NO_2$ requires 234.1489; retention time (achiral analytical HPLC method 1) = 4.9 min.

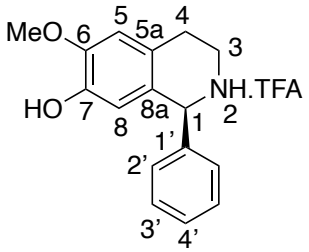
8.7.5 Products from methyltransferase reactions

(1*S*)-1-Cyclohexyl-6-methoxy-1,2,3,4-tetrahydroisoquinolin-7-ol.TFA (171)[149]



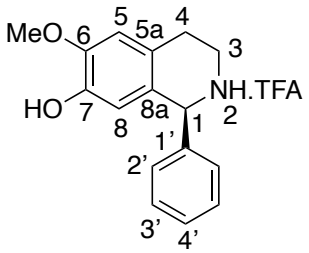
171 was generated using *Mx*SafC as the methyltransferase on a 10 mL, 5 mM scale. Purified by preparative HPLC method 10 and analysed by chiral HPLC method 5. The product was isolated as a white solid (4.4 mg, 25%). $[\alpha]_D^{25}$ 1.1 (*c* 0.022, MeOH); ν_{max}/ cm^{-1} (thin film): 2930, 2855, 1670, 1614, 1516; ^{1}H-NMR (700 MHz; CD_3OD) δ 6.77 (1H, s, 5-H), 6.68 (1H, s, 8-H), 4.28 (1H, d, *J* = 5.2 Hz, 1-H), 3.84 (1H, s, OC*H*$_3$), 3.56 – 3.51 (1H, m, 3-H), 3.30 – 3.26 (1H, m, 3-H), 3.06 – 2.99 (1H, m, 4-*H*H), 2.93 (1H, dt, *J* = 16.8, 5.3 Hz, 4-H*H*), 2.08 – 2.02 (1H, m, 1'-H), 1.89 – 1.83 (1H, m), 1.83 – 1.71 (3H, m), 1.59 – 1.53 (1H, m), 1.42 – 1.34 (1H, m), 1.34 – 1.26 (1H, m), 1.22 (2H, qd, *J* = 12.6, 3.4 Hz); ^{13}C-NMR (176 MHz; CD_3OD) δ 149.0 (C-6 or C-7), 146.8 (C-6 or C-7), 124.3 (C-5a or C-8a), 123.0 (C-5a or C-8a), 114.1 (C-8), 112.6 (C-5), 61.3 (C-1), 56.3 (O**C**H$_3$), 41.9 (C-1'), 41.5 (C-3). 30.9 (C-2', C-3' or C-4'), 27.6, (C-2', C-3' or C-4') 27.2 (C-2', C-3' or C-4'), 27.0 (C-2', C-3' or C-4'), 25.7 (C-4); *m/z* [ES+] 262 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 262.1802; $C_{16}H_{23}NO_2$ requires 262.1802; retention time (achiral analytical HPLC method 1): 6.2 min.

(1*S*)-6-Methoxy-1-phenyl-1,2,3,4-tetrahydroisoquinolin-7-ol.TFA (172)[149]



C6-OMe-172 was generated using *Rn*COMT as the methyltransferase on a 5 mL, 10 mM scale. 87:13 ratio of **C6-OMe-172**:**C7-OMe-174** was generated. Purified by preparative HPLC method 10 analysed by chiral HPLC method 5. The product was isolated as a white solid (3.9 mg, 22%). ν_{max}/ cm^{-1} (thin film): 2932, 2855, 2801, 2582, 1664, 1514; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.51 – 7.45 (3H, m, 3'-H and 4'-H), 7.38 – 7.33 (2H, m, 2'-H), 6.84 (0.9H, s, 5-H), 6.74 (0.1H, s, 5-H), 6.31 (0.1H, s, 8-H), 6.21 (0.9H, s, 8-H), 5.66 (0.1H, s, 1-H), 5.60 (0.9H, s, 1-H), 3.87 (2.7H, s, OCH_3), 3.62 (0.3H, s, OCH_3), 3.53 – 3.39 (2H, m, 3-H), 3.24 – 3.01 (2H, m, 4-H); ^{13}C-NMR (176 MHz, CD_3OD) δ 149.5 (C-6 or C-7), 146.9 (C-6 or C-7), 137.8 (C-1'), 130.9 (C-2' or C-3'), 130.8 (C-2' or C-3'), 130.3 (C-4'), 124.4 (C-5a or C-8a), 124.2 (C-5a or C-8a), 115.1 (C-8), 112.3 (C-5), 60.7 (C-1), 56.3 (O*C*H$_3$), 41.2 (C-3), 25.7 (C-4); *m/z* [ES+] 256 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 256.1332; $C_{16}H_{18}NO_2$ requires 256.1332; retention time (achiral analytical HPLC method 1): 5.7 min.

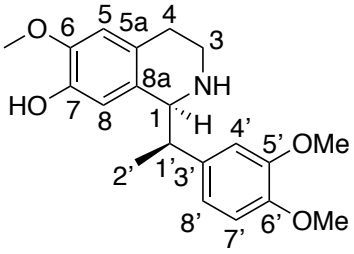
(1*S*)-6-Methoxy-1-phenyl-1,2,3,4-tetrahydroisoquinolin-7-ol.TFA (172)[149]



(157) was generated using *Mx*SafC as the methyltransferase on a 5 mL, 10 mM scale. Purified by preparative HPLC method 10 analysed by chiral HPLC method 5. The product was isolated as a white solid (2.7 mg, 15%). $[\alpha]_D^{25}$ +1.35 (*c* 0.20, MeOH)**;** ν_{max}/

cm^{-1} (thin film): 2932, 2855, 2801, 2582, 1664, 1514; ^{1}H-NMR (700 MHz; CD_3OD) δ 7.50 – 7.45 (3H, m, 3'-H and 4'-H), 7.38 – 7.32 (2H, m, 2'-H), 6.85 (1H, s, 5-H), 6.21 (1H, s, 8-H), 5.60 (1H, s, 1-H), 3.87 (3H, s, OCH_3), 3.52 – 3.47 (1H, m, 3-*H*H), 3.47 – 3.41 (1H, m, 3-H*H*), 3.23 - 3.17 (1H, m, 4-*H*H), 3.07 (1H, dt, *J* = 17.2, 5.5 Hz, 4-H*H*); ^{13}C-NMR (176 MHz; CD_3OD) δ 149.5 (C-6 or C-7), 146.9 (C-6 or C-7), 137.8 (C-1'), 130.9 (C-2' or C-3'), 130.8 (C-2' or C-3'), 130.3 (C-4'), 124.4 (C-5a or C-8a), 124.2 (C-5a or C-8a), 115.1 (C-8), 112.3 (C-5), 60.7 (C-1), 56.3 (O*C*H$_3$), 41.2 (C-3), 25.7 (C-4); *m/z* [ES+] 256 ([M+H]$^+$, 100%), 257); *m/z* [HRMS ES+] found [M+H]$^+$ 256.1332; $C_{16}H_{18}NO_2$ requires 256.1332; retention time (achiral analytical HPLC method 1): 5.7 min.[149]

(*S*)-1-((*R*)-1-(3,4-Dimethoxyphenyl)ethyl)-6-methoxy-1,2,3,4-tetrahydroisoquinolin-7-ol (125)



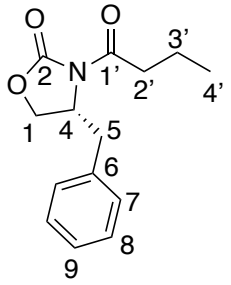
An NCS-mediated reaction between dopamine HCl and 2-(3,4-dimethoxyphenyl)propanal was performed as detailed in the standard procedure, followed by lyophilisation. The residue was resuspended in buffer (4 mL, 250 mM HEPES, 200 mM $MgCl_2$, 2 M KCl, pH 7.5). ATP (2 mL, 100 mM in dH_2O), L-methionine (2 mL, 100 mM in dH_2O), *Ec*MAT (888 μL, at 9 mg mL^{-1} purified in 50 mM HEPES, pH 7.5, 0.4 mg mL^{-1} final concentration), *Ec*MTAN (200 μL, 2.5 mg mL^{-1} purified, 0.025 mg mL^{-1} final concentration) and *Rn*COMT (10% *v/v* lysate) were added. The reaction volume was adjusted to 20 mL by addition of HEPES buffer (50 mM, pH 7.5). Reactions were performed at 37 °C for 24 h. The resulting residue was lyophilised for a further reaction step or purified by preparative HPLC (method 9) to give the product as a white solid (5.3 mg, 12% isolated yield). $[\alpha]_D^{25}$ +4.0 (*c* 0.45, MeOH)**;** ν_{max}/ cm^{-1} (thin film):

2921, 2851, 1710, 1688, 1513; ^{1}H-NMR (700 MHz; CD_3OD) δ 6.95 (1H, d, *J* = 8.3 Hz, 7'-H), 6.83 (1H, dd, *J* = 8.1, 2.1 Hz, 8'-H), 6.80 (1H, d, *J* = 1.8 Hz, 4'-H), 6.76 (1H, s, 5-H), 6.50 (1H, s, 8-H), 4.62 (1H, d, *J* = 6.0 Hz, 1-H), 3.84 (3H, s, OC*H*$_3$), 3.82 (3H, s, OC*H*$_3$), 3.78 (3H, s, OC*H*$_3$), 3.56 – 3.51 (1H, m, 1'-H), 3.46 – 3.41 (1H, m, 3-*H*H), 3.23 – 3.17 (1H, m, 3-H*H*), 3.01 – 2.90 (2H, m, 4-H). 1.30 (2H, d, *J* = 7.2 Hz, 2'-H); ^{13}C-NMR (176 MHz; CD_3OD) δ 150.7 (C-5' or C-6'), 150.0 (C-5' or C-6'), 149.1 (C-6), 146.6 (C-7), 134.4 (C-3'), 124.2 (C-5a or C-8a), 123.9 (C-5a or C-8a), 121.4 (C-8'), 114.5 (C-8), 113.1 (C-4' or C-7'), 113.1 (C-4' or C-7'), 112.4 (C-5), 61.8 (C-1), 56.4 (O*C*H$_3$), 56.4 (O*C*H$_3$), 56.3 (O*C*H$_3$), 43.0 (C-1'), 41.1 (C-3), 25.6 (C-4), 15.2 (C-2'); *m/z* [ES+] 344 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 344.1853; $C_{20}H_{26}NO_4$ requires 344.1862; retention time (achiral analytical HPLC method 2): 8.9 min.

8.8 Small molecule syntheses

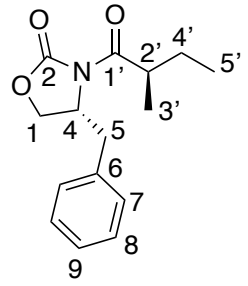
8.8.1 Aldehyde syntheses

(*R*)-4-Benzyl-3-butyryloxazolidin-2-one (105)[148,167]



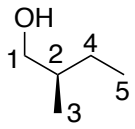
Under anhydrous conditions, a solution of (*R*)-4-benzyloxazolidin-2-one (5.00 g, 28.2 mmol) in THF (40 mL) was prepared and cooled to -78 °C. *n*-Butyllithium (13.5 mL, 31.0 mmol, 2.3 M in cyclohexanes) was added dropwise over 30 min and the reaction mixture stirred for 30 min. Butryl chloride (3.50 mL, 33.8 mmol) was added and the reaction mixture stirred for 1 h and then warmed to room temperature and stirred for 1 h. The reaction was quenched by addition of saturated aqueous NH_4Cl (25 mL) and the two phases separated. The aqueous phase was extracted with dichloromethane (3 x 25 mL). The organic layers were combined, washed with 1 M NaOH (20 mL) and brine (20 mL), dried (Na_2SO_4) and concentrated under reduced pressure. The resulting residue was purified by column chromatography (2.5% ethyl acetate in petroleum ether 40-60). This gave the product as a pale-yellow oil (6.73 g, 96.2%). ^{1}H-NMR (700 MHz; $CDCl_3$) δ 7.36 – 7.31 (2H, m, 7-H), 7.29 – 7.23 (1H, m, 9-H), 7.22 – 7.18 (2H, m, 8-H), 4.70 – 4.62 (1H, m, 4-H), 4.20 - 4.13 (2H, m, 5-H), 3.28 (1H, dd, *J* = 13.5, 2.9 Hz, 1-H), 2.97 - 2.83 (2H, m, 2'-H), 2.76 (1H, dd, *J* = 13.5, 3.9 Hz, 1-H), 1.75 – 1.70 (2H, m, 3'-H), 1.00 (3H, t, *J* = 7.4 Hz, 4'-H); ^{13}C-NMR (176 MHz, $CDCl_3$) δ 173.3 (C-1'), 153.6 (C-2), 135.5 (C-6), 129.6, 129.1, 127.4, 66.3 (C-1), 55.2 (C-4), 38.0 (C-5), 37.5 (C-2'), 17.8 (*C-3'*), 13.8 (C-4'); *m/z* [ES$^+$] 248 ([M+H]$^+$, 100%).

(*R*)-4-Benzyl-3-((*R*)-2-methylbutanoyl)oxazolidin-2-one (106)[148,167]



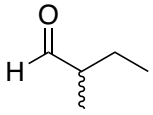
A solution of (*R*)-4-benzyl-3-butyryloxazolidin-2-one (6.60 g, 26.7 mmol) in THF (40 mL) was prepared under anhydrous conditions and cooled to -78 °C. NaHMDS (2 M in THF, 23.6 mL, 47.3 mmol) was added dropwise over 5 minutes and the reaction mixture was stirred (*ca.* 30 min). Methyl iodide (9.84 mL, 158 mmol) was added, and the reaction mixture was stirred until completion (*ca.* 1 h). The reaction was quenched by addition of brine (100 mL) and dichloromethane (150 mL) and warmed to room temperature. The phases were separated, and the aqueous phase diluted with water (100 mL) and extracted with dichloromethane (3 x 75 mL). The organic phases were combined, dried (Na_2SO_4) and concentrated under reduced pressure. The resulting oil was purified by column chromatography (3% ethyl acetate in petroleum ether 40-60) to give the product as a pale-yellow oil (5.60 g, 81.0%). ^{1}H-NMR (700 MHz; $CDCl_3$) δ 7.35 – 7.31 (2H, m, 8-H), 7.26 – 7.24 (1H, m, 9-H), 7.23 – 7.19 (2H, m, 7-H), 4.70 – 4.62 (1H, m, 4-H), 4.22 - 4.15 (2H, m, 5-H), 3.64 (1H, m, 2'-H), 3.27 (1H, dd, *J* = 13.3, 3.4 Hz, 1-H), 2.77 (1H, dd, *J* = 13.3, 3.6 Hz, 1-H), 1.74 – 1.68 (1H, m, 4'-H), 1.48 (1H, m, 4'-H), 1.22 (3H, d, *J* = 7.1 Hz, 3'-H), 0.93 (3H, t, *J* = 7.1 Hz, 5'-H); ^{13}C-NMR (176 MHz; $CDCl_3$) δ 177.3 (C-1'), 153.2 (C-2), 135.5 (C-6), 129.6, 129.1, 127.5, 66.2 (C-1), 55.5 (C-4), 39.3 (C-5), 38.0 (C-2'), 26.5 (C-4'), 17.0 (C-3'), 11.8 (C-5'); *m/z* [ES+] 284 ([M+Na]$^+$, 100%).

(*R*)-2-Methybutan-1-ol ((*R*)-103)[148,168]



Under anhydrous conditions, a solution of (*R*)-4-benzyl-3-((*R*)-2-methylbutanoyl)oxazolidin-2-one (5.56 g, 2.13 mmol) in diethyl ether (70 mL) was prepared and cooled to 0 °C. Methanol (1.1 mL) and lithium borohydride (43.2 mmol, 0.94 g) were added. The reaction mixture was stirred for 10 min then quenched by the dropwise addition of saturated aqueous $NaHCO_3$ (70 mL) and warmed to room temperature. Diethyl ether (70 mL) was added, the phases separated, and the organic phases washed with saturated aqueous $NaHCO_3$ (2 x 50 mL). The organic phases were combined, dried, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (25% ethyl acetate in petroleum ether 40-60) to give the product as a colourless oil (0.680 g, 36.0%) ^{1}H-NMR (700 MHz; $CDCl_3$) δ 3.52 (1H, m, 1-H) 3.43 (1H, m, 1-H), 1.54 (1H, m, 4-H), 1.45 (1H, m, 4-H), 1.15 (1H, m, 2-H), 0.92 (3H, d, J = 7.1 Hz, 3-H), 0.91 (3H, t, J = 7.1 Hz, 5-H); ^{13}C-NMR (176 MHz; $CDCl_3$) δ 68.2 (C-1), 37.5 (C-2), 25.9 (C-4), 16.2 (C-3), 11.5 (C-5); GCMS analysis given in the Appendix.

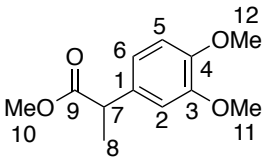
(*R*)/(*S*)-2-Methybutanal (91)[169]



A solution of oxalyl chloride (2 M in dichloromethane, 1.40 mL, 2.79 mmol) in dichloromethane (10 mL) was prepared and cooled to -78 °C. DMSO (295 µL, 5.58 mmol) was added over 10 min and the reaction mixture stirred for 30 min. 2-methylbutan-1-ol (200 µL, 1.86 mmol) in dichloromethane (5 mL) was added, and the reaction mixture stirred for 30 min. Triethylamine (1.55 mL, 11.2 mmol) was added and the reaction mixture stirred for 30 min at -30 °C followed by 30 min at rt. The mixture was diluted with pentane (10 mL) and washed with $NaHCO_3$ (2 x 10 mL) and water (2 x 10 mL). The organic phase was dried and concentrated under reduced pressure to give an oil which was partially purified by silica chromatography (100% dichloromethane) to give the crude aldehyde as a yellow oil. The crude aldehyde ((*S*)=

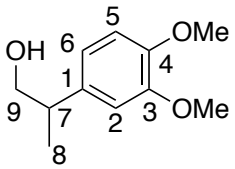
126 mg, 77%, (*R*)=114 mg, 70%) was used directly for enzymatic reactions without further purification. The enantiopurity of the resultant aldehydes was confirmed by GC analysis.

Methyl 2-(3,4-dimethoxyphenyl)propanoate (121)[199]



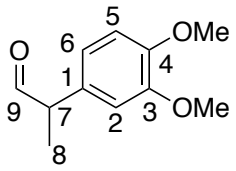
Under anhydrous conditions, a solution of methyl 2-(3,4-dimethoxyphenyl)acetate (2.00 g, 9.52 mmol) in THF (10 mL) was prepared and cooled to -78 °C. A solution of LDA (7.2 mL, 2.0 M in THF, 14.3 mmol) was added dropwise and the reaction mixture stirred (*ca.* 30 min). Iodomethane (1.2 mL, 19 mmol) was added, and the reaction stirred for 1 h, then quenched by the addition of 2 M HCl. The resulting solution was extracted with ethyl acetate (3 x 20 mL), washed with brine (2 x 20 mL), dried ($MgSO_4$) and concentrated under reduced pressure. The resulting residue was purified by flash chromatography (10% ethyl acetate in petroleum ether 40-60) to give the product as a colourless oil (1.35 g, 63%); ^{1}H-NMR (700 MHz; $CDCl_3$): δ 6.86 - 6.80 (3H, m, 2-H, 5-H and 6-H), 3.88 (3H, s, 11-H or 12-H), 3.86 (3H, s, 11-H or 12-H), 3.66 (1H, q, J = 7.3 Hz, 7-H), 3.66 (3H, s, 10-H), 1.48 (3H, d, J = 7.3 Hz, 8-H); ^{13}C-NMR (176 MHz; $CDCl_3$): δ 175.3 (C-9), 149.1 (C-3), 148.3 (C-4), 133.2 (C-1), 119.7 (C-6), 111.3 (C-2), 110.8 (C-5), 56.0 (C-11 or C-12), 56.0 (C-11 or C-12), 52.1 (C-10), 45.1 (C-7), 18.8 (C-8); *m/z* (GC-MS): 224 [M+].

2-(3,4-Dimethoxyphenyl)propan-1-ol (122)[199]



Under anhydrous conditions, a solution of 2-(3,4-dimethoxyphenyl)acetate (858 mg, 3.83 mmol) in THF (15 mL) was prepared, cooled to -78 °C and DIBAL-H (1.0 M in THF, 7.65 mmol, 7.65 mL) was added dropwise over 10 min. The reaction mixture was stirred for 2 h, then quenched by addition of Rochelle's solution (10 mL). The reaction was warmed to 0 °C and stirred for 1 h and then poured into water (10 mL). The reaction mixture was extracted with ethyl acetate (3 x 15 mL), dried with $MgSO_4$ and concentrated under reduced pressure to give the product as a yellow oil (710 mg, 95%); ^{1}H-NMR (600 MHz; $CDCl_3$): δ 6.85 – 6.74 (3H, m, 2-H, 5-H and 6-H), 3.88 (3H, s, 3-OC*H*$_3$ or 4-OC*H*$_3$), 3.86 (3H, s, 3-OC*H*$_3$ or 4-OC*H*$_3$), 3.70 - 3.63 (2H, m, 9-H), 2.85 – 2.92 (1H, m, J = 6.8 Hz, 7-H), 1.25 (3H, d, J = 6.8 Hz, 8-H); ^{13}C-NMR (151 MHz; $CDCl_3$): δ 149.2 (C-3 or C-4), 147.9 (C-3 or C-4), 136.3 (C-1), 119.4 (C-6), 111.5 (C-2 or C-5), 110.9 (C-2 or C-5), 68.9 (C-9), 56.1 (3-O*C*H$_3$ or 4-O*C*H$_3$), 56.0 (3-O*C*H$_3$ or 4-O*C*H$_3$), 42.2 (C-7), 17.9 (C-8); *m/z* [HRMS ES+] found $[M+Na]^+$ 219.0093; $C_{11}H_{16}O_3Na$ requires 219.0092.[199]

2-(3,4-Dimethoxyphenyl)propanal (118)

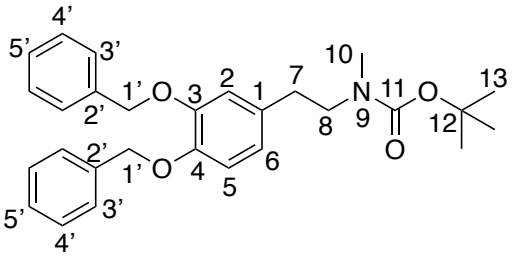


Under anhydrous conditions, a solution of oxalyl chloride (2 M in CH_2Cl_2; 1.92 mL, 3.83 mmol) in dichloromethane (10 mL) was prepared at -78 °C. DMSO (403 μL, 7.64 mmol) was added dropwise and the reaction mixture stirred for 30 min. 2-(3,4-Dimethoxyphenyl)propan-1-ol (500 mg, 2.60 mmol) in dichloromethane (*ca.* 5 mL) was then added and the reaction mixture stirred for 30 min, followed by the addition of triethylamine (2.12 mL, 15.3 mmol) and stirring for 30 min at -30 °C, then 30 min at 0 °C. The mixture was then diluted with pentane (15 mL), washed with sat. aq. $NaHCO_3$ (2 x 20 mL) and water (2 x 20 mL), dried with Na_2SO_4 and concentrated under reduced pressure to give the product as an orange oil (361 mg, 47%). ^{1}H-NMR (700 MHz;

$CDCl_3$): δ 9.66 (1H, d, *J* = 1.5 Hz, 9-H), 6.90 – 6.86 (1H, m, Ar*H*), 6.79 – 6.75 (1H, m, Ar*H*), 6.67 – 6.71 (1H, m, Ar*H*), 3.88 (6H, s, 3-OCH_3 and 4-OCH_3), 3.88 (1H, m, 7-H), 1.43 (3H, t, *J* = 6.9 Hz, 8-H); ^{13}C-NMR (176 MHz; $CDCl_3$): δ 201.2 (C-9), 149.6 (C-3 or C-4), 148.7 (C-3 or C-4), 130.1 (C-1), 120.6 (C-6), 111.8 (C-2 or C-5), 111.5 (C-2 or C-5), 56.1 (3-OCH_3 or 4-OCH_3), 56.1 (3-OCH_3 or 4-OCH_3), 52.7 (C-7), 14.7 (C-8); *m/z* [ES+] 195 ($[M+H]^+$, 100%).

8.8.2 Amine syntheses

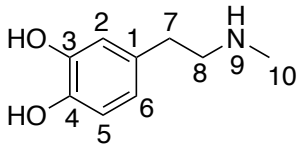
***tert*-Butyl (3,4-bis(benzyloxy)phenethyl)(methyl)carbamate (199)**[270]



A solution of *tert*-butyl (3,4-bis(benzyloxy)phenethyl)carbamate (3.50 g, 8.07 mmol) in DMF (50 mL) was prepared under anhydrous conditions and cooled to 0 °C. The reaction mixture was stirred and sodium hydride (60% in mineral oil; 484 mg, 12.1 mmol) was added, stirred for 1 h, followed by the addition of iodomethane (603 μL, 9.68 mmol). The solution was allowed to warm to room temperature and stirred for 18 h, then diluted with diethyl ether (30 mL) and quenched with H_2O (5 mL). The reaction mixture was extracted with diethyl ether (3 x 30 mL), the organic layers washed with water (3 x 50 mL) and brine (3 x 50 mL) then dried (Na_2SO_4) and concentrated under reduced pressure. The resulting residue was purified by column chromatography (5 – 20% ethyl acetate in petroleum ether 40-60) to give the product as a colourless oil (3.16 g, 90%). Two rotamers of the product were observed by ^{1}H-NMR due to the presence of a tertiary amide. ν_{max}/ cm^{-1} (thin film): 3011, 2929, 1631; ^{1}H-NMR (700 MHz; DMSO-d_6) δ 7.47 - 7.26 (10H, m, Ar*H*), 6.96 – 6.93 (1H, m, Ar*H*), 6.91 – 6.88 (1H, m, Ar*H*), 5.14 – 5.04 (4H, m, PhCH_2O), 2.70 (3H, s, NCH_3), 2.68 – 2.60 (2H, m, CH_2), 2.51 – 2.46 (2H, m, CH_2), 1.42-1.24 (9H, m, (CH_3)$_3$); ^{13}C-NMR (150 MHz; DMSO-d_6) δ 155.7 (C-11), 149.0

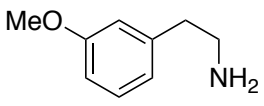
(C-3 or C-4), 147.6 (C-3 or C-4), 137.6 (C-2'), 137.6 (C-2'), 132.9 (C-1), 128.6, 128.6, 127.9. 127.9, 127.5, 127.5, 121.9 (C-6). 116.0 (C-2), 115.5 (C-5), 79.4 (C-12), 71.6 (C-1'), 71.6 (C-1'), 51.1 (C-8), 34.6 (C-10), 34.2 (C-7), 28.6 (C-13); *m/z* [ES+] 448 ($[M+H]^+$, 100%); *m/z* [HRMS ES+] found $[M+H]^+$ 470.2302, $C_{28}H_{33}NO_4$ requires 470.2302.

4-(2-(Methylamino)ethyl)benzene-1,2-diol (195)[263]



A solution of *tert*-butyl (3,4-bis(benzyloxy)phenethyl)(methyl)carbamate (2.22 g, 4.96 mmol) in MeOH (25% *v/v* 37% HCl) was prepared and heated under reflux for 18 h. Solvents were removed under reduced pressure. The resulting residue was resuspended in acetone and the product was isolated by filtration as a white solid (580 mg, 70%). ^{1}H-NMR (600 MHz; D_2O) δ 6.98 (1H, m, 5-H), 6.92 (1H, m, 2-H), 6.83 (1H, m, 6-H), 3.34 (2H, t, *J* = 7.2 Hz, 8-H), 2.98 (2H, t, *J* = 7.2 Hz, 7-H), 2.78 (3H, s, 10-H); ^{13}C-NMR (150 MHz; D_2O) δ 144.9 (C-3), 143.7 (C-4), 129.6 (C-1), 121.8 (C-6), 117.2 (C-2 or C-5), 117.1 (C-2 or C-5), 50.8 (C-8), 33.4 (C-7), 31.6 (C-10); *m/z* [ES+] 168 ($[M+H]^+$, 100%).

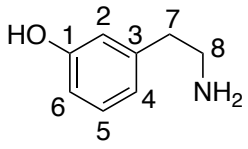
2-(3-Methoxyphenyl)ethan-1-amine (185)[22]



A solution of 3-methyoxyphenylacetonitrile (1.00 g, 6.70 mmol) in THF (150 mL) was prepared under anhydrous conditions. Boron trifluoride diethyl etherate (1 M in THF, 495 μL, 4.02 mmol) was added and the solution stirred under reflux for 10 min. Borane dimethyl sulfide complex (2 M in THF, 20.1 mL, 20.1 mmol) was added and the reaction stirred under reflux for 3 h. The reaction was cooled to 0 °C, 1 M HCl (100 mL) was

added and the reaction stirred for 1 h, then warmed to room temperature and stirred for 18 h. The pH was adjusted to 13 and the product extracted into dichloromethane (3 x 100 mL). The organic phases were dried with anhydrous $NaSO_4$ and concentrated under reduced pressure to give the product as a white solid (0.99 g, 98%). The product was taken through to the next step without further purification.

3-(2-Aminoethyl)phenol (51)[22]



Under anhydrous conditions, a solution of 2-(3-methoxyphenyl)ethan-1-amine (0.95 g, 6.3 mmol) in dichloromethane (20 mL) was prepared and cooled to 0 °C. Boron tribromide (1 M in dichloromethane, 15.7 mL, 15.7 mmol) was added dropwise and the reaction warmed to room temperature. The reaction was then heated under reflux for 16 h, quenched by the addition of methanol (5 mL) and concentrated under reduced pressure. The residue was resuspended in ethyl acetate (80 mL), poured into ice and 1 M HCl (100 mL) and stirred for 30 min. The two phases were separated, and the aqueous layer was concentrated under reduced pressure to give the product as a pale brown solid. The product was recrystallised from *n*-hexane to give a white solid (0.88 g, 81%). ^{1}H-NMR (700 MHz; CD_3OD) δ 7.26 (1H, t, J = 7.9 Hz, 5-H), 6.88 – 6.76 (3H, m, 2-H, 4-H, 6-H), 3.23 (2H, t, J = 7.0 Hz, H-8), 2.91 (2H, t, J = 7.0 Hz, H-7); ^{13}C-NMR (176 MHz; CD_3OD) δ 156.4 (C-1), 139.1 (C-3), 131.0 (C-5), 121.5 (C-2), 116.3 (C-4), 114.8 (C-6), 41.0 (C-8), 33.2 (C-7); *m/z* [ES^+] 138 ($[M+H]^+$, 100%).

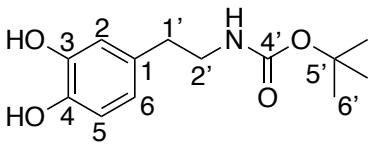
8.8.3 Generation of racemic THIQ using a biomimetic phosphate reaction[15]

The phenethylamine (10 mM), sodium ascorbate (10 mM) and aldehyde (20 mM) in KPi buffer (300 mM, pH 6) with 50% *v/v* MeCN was prepared and heated to 60 °C for 18 h. Reactions were performed on either an analytical (200 μL) or preparative (10-20 mL)

scale and products isolated in an identical manner to the corresponding products generated by enzyme-catalysed reactions.

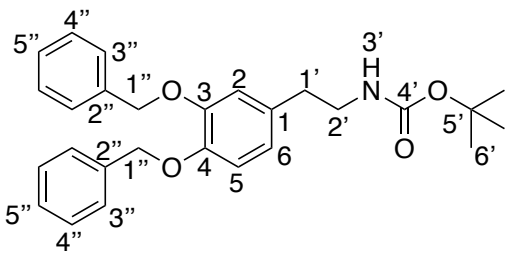
8.8.4 Synthesis of reaction mimics for co-crystallisation

tert-Butyl (3,4-dihydroxyphenethyl)carbamate (66)[30]



A solution of dopamine hydrochloride (3.00 g, 15.8 mmol) in THF (10 mL) and saturated aqueous $NaHCO_3$ (6 mL) was prepared and stirred. Boc anhydride (3.75 g, 17.2 mmol) was added and the reaction mixture stirred for 2 h at rt. The aqueous fraction was washed with ethyl acetate (3 x 30 mL) and the organic fractions combined and washed with brine (3 x 20 mL). The organic phase was dried with $MgSO_4$ and concentrated under reduced pressure to give the desired product as a white solid (3.9 g, 89%).[251] ^{1}H-NMR (400 MHz; DMSO-d_6) δ 8.75 (1H, s, O*H*), 8.65 (1H, s, O*H*), 6.85 – 6.81 (1H, m, N*H*), 6.62 – 6.58 (1H, m, 5-H), 6.56 – 6.52 (1H, m, 2-H), 6.42 – 6.39 (1H, m, 6-H), 3.07 – 3.00 (2H, m, 2'-H), 2.51 – 2.46 (2H, m, 1'-H), 1.37 (9H, s, 6'-H); ^{13}C-NMR (151 MHz; DMSO-d_6) δ 155.9 (C-4), 144.9 (C-3), 143.4 (C-4), 130.2 (C-1), 119.1 (C-6), 115.8 (C-2), 115.3 (C-5), 77.4 (C-5'), 42.1 (C-2'), 35.0 (C-1'), 28.2 (C-6'); *m/z* [ES+] 254 ($[M+H]^+$, 100%).[30]

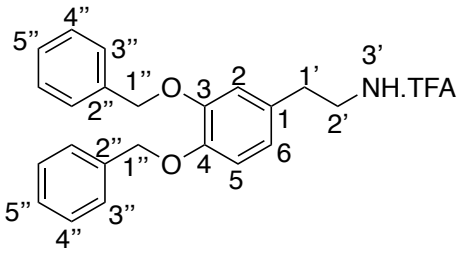
tert-Butyl (3,4-bis(benzyloxy)phenethyl)carbamate (67)[30]



Under an inert atmosphere, a solution of *tert*-butyl (3,4-dihydroxyphenethyl)carbamate (3.90 g, 15.4 mmol) and K_2CO_3 (12.4 g, 90.0 mmol) in anhydrous DMF was prepared

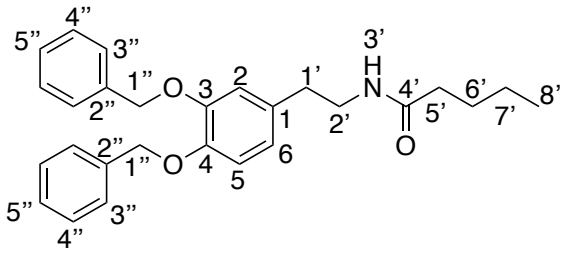
and stirred at room temperature (*ca.* 30 min). Benzyl bromide (5.51 mL, 46.4 mmol) was added dropwise at 0 °C and the reaction mixture stirred (*ca*. 16 h). The reaction mixture was filtered through Celite and the solid washed with diethyl ether (3 x 40 mL). This resulted in the precipitation of the product which was dried under vacuum to give a colourless solid (95% crude yield).[241] ^{1}H-NMR (400 MHz; $CDCl_3$) δ 7.46 - 7.27 (10H, m, 3''-H, 4''-H and 5''-H), 6.89 – 6.83 (1H, m, 5-H), 6.74 (1H, s, 2-H), 6.71 – 6.67 (1H, m, 6-H), 5.17 – 5.11 (4H, m, 1''-H), 3.33 – 3.29 (2H, m, 2'-H), 2.70 – 2.66 (2H, m, 1'-H), 1.45 – 1.40 (9H, s, 6'-H); ^{13}C-NMR (151 MHz; $CDCl_3$) δ 156.0 (C-4'), 149.0 (C-3), 147.7 (C-4), 137.5 (C-2''), 137.6 (C-2''), 132.5 (C-1), 128.6, 127.9, 127.9, 127.5, 127.4, 121.8 (C-6), 115.9 (C-2), 115.4 (C-5), 79.3 (C-5'), 71.5 (C-1''), 71.5 (C-1''), 41.9 (C-2'), 36.6 (C-1'), 28.5 (C-6'); *m/z* [ES+] 378 ([M+H-^{t}Bu]$^{+}$, 100%).[30]

3,4-Di(benzyloxy)phenethylamine (68)[30]



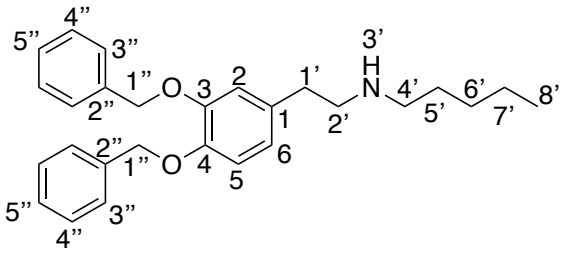
A solution of *tert*-butyl (3,4-bis(benzyloxy)phenethyl)carbamate (0.30 g, 0.70 mmol) in dichloromethane (8 mL) and TFA (2 mL) was prepared and stirred at room temperature (*ca.* 30 min). Solvents were removed under reduced pressure to give the product as a colourless oil (0.23 g, 100%).[241] ^{1}H-NMR (400 MHz; $CDCl_3$) δ 7.44-7.27 (10H, m, 2''-H to 5''-H), 6.92 – 6.89 (1H, m, 5-H), 6.75 – 6.72 (1H, m, 2-H), 6.69 – 6.64 (1H, m, 6-H), 5.12 – 5.08 (4H, m, 1''-H), 3.08 – 3.02 (2H, m, 2'-H), 2.75 (2H, t, *J* = 6.9 Hz, 1'-H); ^{13}C-NMR (600 MHz; $CDCl_3$) δ 148.8 (C-3), 148.2 (C-4), 129.4 (C-1), 129.1-127.7 (C-3'' to C-5')', 122.2 (C-6), 115.8 (C-2), 115.4 (C-5), 71.8 (C-1')', 71.5 (C-1''), 41.6 (C-2), 37.8 (C-1'); *m/z* [ES+] 334 ([M+H]$^{+}$, 100%).

***N*-(3,4-Bis(benzyloxy)phenethyl)pethanamide (69)**[30]



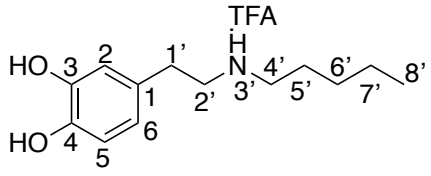
A solution of 3,4-di(benzyloxy)phenethylamine (230 mg, 0.70 mmol) and triethylamine (290 μL, 0.75 mmol) in dichloromethane (8 mL) was prepared under anhydrous conditions and stirred. Pentanoic acid (82 μL, 0.75 mmol) and thionyl chloride (54 μL, 0.75 mmol) were added and the reaction mixture stirred at room temperature for 30 min. Solvents were removed under reduced pressure and the residue was re-suspended in dichloromethane (10 mL) washed with 1 M HCl (3 x 10 mL) and 1 M NaOH (3 x 10 mL), dried (Na_2SO_4) and concentrated under reduced pressure. The resulting white solid was purified by silica chromatography (15-50% ethyl acetate in petroleum ether 40-60) to give the product (95 mg, 38%). ν_{max}/ cm^{-1} (thin film): 3305, 2966, 2926, 2869, 1639, 1598, 1545, 1511; ^{1}H-NMR (600 MHz; $CDCl_3$) δ 7.47 – 7.43 (4H, m, 3’’-H), 7.39 – 7.34 (4H, m, 4’’-H), 7.33 – 7.29 (2H, m, 5’’-H), 6.80 – 6.86 (1H, m, 5-H), 6.79 – 6.74 (1H, m, 2-H), 6.72 – 6.69 (1H, m, 6-H), 5.18 – 5.12 (4H, m, 1’’-H), 3.45 (2H, q, *J* = 6.6 Hz, 2’-H), 2.70 (2H, t, *J* = 7.2 Hz, 1’-H), 2.08 (2H, t, *J* = 7.8 Hz, 5’-H), 1.56 (2H, m, 6’-H), 1.30 (2H, m, 7’-H), 0.89 (3H, t, *J* = 7.8 Hz, 8’-H); ^{13}C-NMR (151 MHz; $CDCl_3$) δ 173.2 (C-4’), 149.0 (C-3), 147.8 (C-4), 137.5 (C-2’’), 137.3 (C-2’’), 132.3 (C-1), 128.6 (C-4’’), 128.6 (C-4’’), 128.0 (C-5’’), 127.9 (C-5’’), 127.4 (C-3’’), 127.4 (C-3’’), 121.7 (C-6), 115.7 (C-2), 115.4 (C-5), 40.6 (C-2’), 36.7 (C-1’), 35.3 (C-5’), 27.9 (C-6’), 22.5 (C-7’), 13.9 (C-8’); *m/z* [ES+] 418 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 418.2394; $C_{27}H_{32}NO_3$ requires 418.2382.

***N*-(3,4-Bis(benzyloxy)phenethyl)pentan-1-amine (71)**[30]



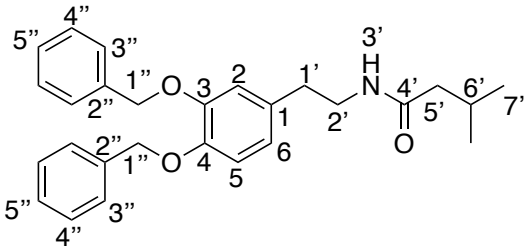
Under anhydrous conditions, a solution of *N*-(3,4-bis(benzyloxy)phenethyl) pethanamide (90 mg, 0.21 mmol) in THF (6.5 mL) was prepared. Boron trifluoride etherate (14 μL, 0.11 mmol) was added and the reaction mixture heated at reflux for 10 min. Boron dimethyl sulfide complex (290 μL, 0.64 mmol) was added and the reaction mixture stirred under reflux for 3 h. The solution was then cooled to 0 °C and 10 % HCl (3.5 mL) was added, and the reaction mixture stirred at 0 °C for 1 h then at rt for 2 h. The pH of the solution was adjusted to 13 by addition of 2 M NaOH and the product extracted into dichloromethane (3 x 8 mL). The organic phase was dried with Na_2SO_4 and concentrated to give the desired product as a colourless oil (73 mg, 86%). ν_{max}/ cm^{-1} (thin film): 2920, 2850, 1604, 1586, 1508; ^{1}H-NMR (600 MHz; $CDCl_3$) δ 7.48 – 7.42 (4H, m, 3’’-H), 7.39 – 7.35 (4H, m, 4’’-H), 7.32 – 7.28 (4H, m, 5’’-H), 6.89 – 6.84 (1H, m, 5-H), 6.83 – 6.78 (1H, m, 2-H), 6.74 – 6.70 (1H, m, 6-H), 5.18 – 5.12 (4H, m, 1’’-H), 2.80 (2H, t, *J* = 7.1 Hz, 2’-H), 2.71 (2H, t, *J* = 7.1 Hz, 1’-H), 2.57 (2H, t, *J* = 7.4 Hz, 4’-H), 1.47 – 1.42 (2H, m, 5’-H), 1.34 – 1.23 (4H, m, 6’-H and 7’-H), 0.88 (3H, t, *J* = 6.0 Hz, 8’-H); ^{13}C-NMR (600 MHz; $CDCl_3$) δ 149.1 (C-3), 147.6 (C-4), 137.6 (C-2’)’, 137.5 (C-2’’), 129.2 (C-1), 128.6 (C-4’’), 128.3 (C-4’’), 127.9 (C-5’’), 127.9 (C-5’’), 127.5 (C-3’’), 127.5 (C-3’’), 121.7 (C-6’’), 116.0 (C-2), 115.6 (C-5), 71.7 (C-1’’), 71.5 (C-1’’), 51.3 (C-4’), 50.0 (C-2’), 35.9 (C-1’), 29.7 (C-5’), 26.6 (C-6’), 22.7 (C-7’), 14.2 (C-8’); *m/z* [ES+] 404 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 404.2579; $C_{27}H_{34}NO_2$ requires 404.2589.

***N*-(3,4-Dihydroxyphenethyl)pentanamine.TFA (73)**[30]



A solution of *N*-(3,4-bis(benzyloxy)phenethyl)pentan-1-amine (55 mg, 0.14 mmol) in methanol (6 mL) and 37% HCl (1.2 mL) was prepared and stirred under anhydrous conditions for 18 h. Water (5 mL) was added and solvents removed under reduced pressure. The resulting residue was purified by preparative-HPLC (method 8) and isolated as a white solid (TFA salt, 9.0 mg, 21%). ν_{max}/ cm^{-1} (thin film): 3038, 2962, 2862, 1669, 1605, 1529; ^{1}H-NMR (600 MHz; $CDCl_3$) δ 6.75 – 6.72 (1H, m, 5-H), 6.71 – 6.68 (1H, m, 2-H), 6.60 – 6.56 (1H, m, 6-H), 3.17 (2H, t, *J* = 7.8 Hz, 2'-H), 2.98 (2H, t, *J* = 7.8 Hz, 1'-H), 2.82 (2H, t, *J* = 8.0 Hz, 4'-H), 1.69 – 1.65 (2H, m, 5'-H), 1.36 – 1.27 (4H, m, 6'-H and 7'-H), 0.95 (3H, t, *J* = 6.9 Hz, 8'-H); ^{13}C-NMR (600 MHz; $CDCl_3$) δ 146.8 (C-3), 145.7 (C-4), 128.9 (C-1), 120.9 (C-6), 116.7 (C-5), 116.7 (C-20, 50.3 (C-2' or C-4'), 49.6 (C-2' or C-4'), 32.8 (C-1'), 29.6 (C-5'), 26.9 (C-6'), 23.2 (C-7'), 14.1 (C-8'); *m/z* [ES+] 224 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 224.1651; $C_{13}H_{22}NO_2$ requires 224.1645.

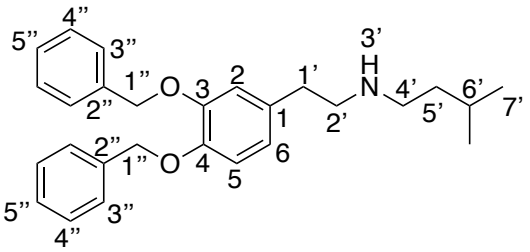
***N*-(3,4-Bis(benzyloxy)phenethyl)-3-methylbutanamide (70)**[30]



A solution of 3,4-di(benzyloxy)phenethylamine (280 mg, 0.61 mmol) and triethylamine (250 μL, 1.8 mmol) in dichloromethane (8 mL) was prepared under anhydrous conditions and stirred. 3-Methylbutanoic acid (82 μL, 0.75 mmol) and thionyl chloride (54 μL, 0.75 mmol) were added and the reaction mixture stirred at room temperature for 30 min. Solvents were removed under reduced pressure and the residue was re-

suspended in dichloromethane (8 mL) washed with 1 M HCl (3 x 10 mL) followed by 1 M NaOH (3 x 10 mL), dried with anhydrous Na_2SO_4 and concentrated under reduced pressure. The resulting white solid was purified by silica chromatography (15-50% ethyl acetate in petroleum ether 40-60) to yield the product as a white solid (130 mg, 50%). ν_{max}/ cm^{-1} (thin film): 3297, 3061, 3030, 2952, 2925, 2865, 1636, 1605, 1586, 1546, 1514; ^{1}H-NMR (600 MHz; $CDCl_3$) δ 7.48 – 7.42 (4H, m, 3''-H), 7.39 – 7.35 (4H, m, 4''-H), 7.34 – 7.29 (4H, m, 5''-H), 6.90 – 6.86 (1H, m, 5-H), 6.82 – 6.77 (1H, m, 2-H), 6.73 – 6.68 (1H, m, 6-H), 5.18 – 5.12 (4H, m, 1''-H), 3.46 (2H, q, *J* = 6.5 Hz, 2'-H), 2.71 (2H, t, *J* = 6.5 Hz, 1'-H), 2.09 – 2.03 (1H, m, 6'-H), 1.94 (2H, d, *J* = 6.5 Hz, 5'-H), 0.91 (6H, d, *J* = 6.5 Hz, 7'-H); ^{13}C NMR (151 MHz; $CDCl_3$) δ 172.5 (C-4'), 149.1 (C-3), 147.8 (C-4), 137.5 (C-2''), 137.4 (C-2''), 132.4 (C-1), 128.6 (C-4''), 128.6 (C-4''), 127.9 (C-5''), 127.5 (C-3''), 121.7 (C-6), 115.9 (C-2), 115.6 (C-5), 71.6 (C-1''), 71.4 (C-1''), 46.3 (C-5'), 40.5 (C-2'), 35.4 (C-1'), 26.2 (C-6'), 22.5 (C-7'); *m/z* [MS ES+] 418 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 418.2407 $C_{27}H_{32}NO_2$ requires 418.2382.

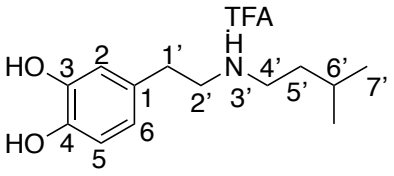
_N_-(3,4-Bis(benzyloxy)phenethyl)-3-methylbutan-1-amine (72)[30]



Under anhydrous conditions, a solution of *N*-(3,4-bis(benzyloxy)phenethyl)-3-methylbutanamide (120 mg, 0.29 mmol) in THF (8.5 mL) was prepared. Boron trifluoride diethyl etherate (19 μL, 0.15 mmol) was added and the reaction heated under reflux for 10 min. Borane dimethyl sulfide complex (2.0 M in THF, 430 μL, 0.87 mmol) was added dropwise and the reaction heated under reflux for a further 3 h. The reaction was cooled to 0 °C, 10% HCl was added (4.5 mL) and the reaction stirred for a further 1 h, followed by 18 h at rt. The pH of the reaction was adjusted to 13 by addition of 2 M NaOH then extracted into dichloromethane (3 x 10 mL). The organic phases were combined, dried

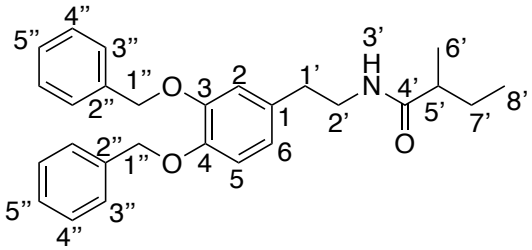
with Na_2SO_4 and concentrated under reduced pressure. The resulting residue was taken through crude to the next reaction step without further purification. *m/z* [HRMS ES+] found $[M+H]^+$ 404.2590; $C_{27}H_{34}NO_2$ requires 404.2589.

_N_-(3,4-Dihydroxyphenethyl)-3-methylbutan-1-amine.TFA (74)[30]



A solution of *N*-(3,4-bis(benzyloxy)phenethyl)-3-methylbutan-1-amine (75 mg, 0.19 mmol) in MeOH (10 mL) and 37% HCl (2 mL) was prepared and heated under reflux for 18 h. Solvents were removed under reduced pressure and the resulting residue was purified by preparative HPLC (method 8) to give the product as a white solid (TFA salt, 32 mg, 53%). ν_{max}/ cm^{-1} (thin film): 2960, 2871, 1667, 1606, 1521; ^{1}H-NMR (600 MHz; $CDCl_3$) δ 6.75 – 6.71 (1H, m, 5-H), 6.70 – 6.67 (1H, m, 2-H), 6.60 – 6.55 (1H, m, 6-H), 3.17 (2H, t, *J* = 7.6 Hz, 2'-H), 3.04 – 2.97 (2H, m, 1'-H), 2.83 (2H, t, *J* = 8.2 Hz, 4'-H), 1.68 – 1.63 (1H, m, 6'-H), 1.57 – 1.52 (2H, m, 5'-H), 0.96 (6H, d, *J* = 6.4 Hz, 7'-H); ^{13}C-NMR (151 MHz; $CDCl_3$) δ 146.8 (C-3), 145.7 (C-4), 128.9 (C-1), 120.9 (C-6), 116.7 (C-5), 116.7 (C-2), 50.3 (C-2' or C-4'), 47.6 (C-2' or C-4'), 35.9 (C-5'), 32.9 (C-1'), 27.1 (C-6'), 22.5 (C-7'); *m/z* [ES+] 223 ($[M+H]^+$, 100%); *m/z* [HRMS ES+] found $[M+H]^+$ 223.1567; $C_{13}H_{21}NO_2$ requires 223.1567.

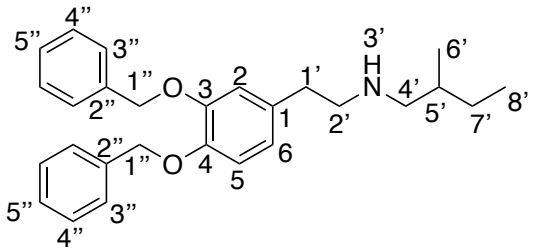
_N_-(3,4-Bis(benzyloxy)phenethyl)-2-methylbutanamide (107)[30]



Under anhydrous conditions, a solution of 3,4-di(benzyloxy)phenethylamine (300 mg, 0.70 mmol) and 2-methylbutyric acid (82 μL, 0.75 mmol) in dichloromethane (8 mL) was

prepared and stirred. Triethylamine (290 μL, 2.1 mmol) was added followed by thionyl chloride (54 μL, 0.75 mmol). The reaction mixture was stirred for 30 min then solvents removed under reduced pressure. The residue was resuspended in dichloromethane (8 mL) and washed with 1 M HCl (3 x 10 mL) and 1 M NaOH (3 x 10 mL). The organic phase was dried with Na_2SO_4 and concentrated under reduced pressure. The resulting residue was purified by column chromatography (20 – 50% ethyl acetate in petroleum ether 40-60) and the pure product isolated as a white solid (69 mg, 24%). ν_{max}/ cm^{-1} (thin film): 3058, 3028, 2948, 2921, 2861, 1603, 1586, 1508; ^{1}H-NMR (600 MHz; $CDCl_3$) δ 7.47 – 7.43 (4H, m, 3''-H), 7.39 – 7.34 (4H, m, 4''-H), 7.33 – 7.29 (4H, m, 5''-H), 6.89 – 6.85 (1H, m, 5-H), 6.83 – 6.79 (1H, m, 2-H), 6.72 – 6.68 (1H, m, 6-H), 5.17 – 5.12 (4H, m, 1''-H), 3.49 – 3.44 (2H, m, 2'-H), 2.75 – 2.69 (2H, m, 1'-H), 2.00 – 1.95 (1H, m, 5'-H), 1.64 – 1.58 (2H, m, 7'-H), 1.08 (3H, d, *J* = 7.2 Hz, 6'-H), 0.85 (3H, t, *J* = 7.2 Hz, 8'-H); ^{13}C NMR (151 MHz; $CDCl_3$) δ 176.4 (C-4'), 149.2 (C-3), 147.8 (C-4), 137.5 (C-2''), 137.4 (C-2''), 132.5 (C-1), 128.6, 128.6, 127.9, 127.9, 127.5, 127.5, 121.8 (C-6), 115.9 (C-2), 115.6 (C-5), 71.6 (C-1''), 71.5 (C-1''), 43.4 (C-5'), 40.5 (C-2'), 35.4 (C-1'), 27.4 (C-7'), 17.6 (C-6'), 12.0 (C-8'); *m/z* [ES+] 418 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 418.2389; $C_{27}H_{32}NO_2$ requires 418.2382.

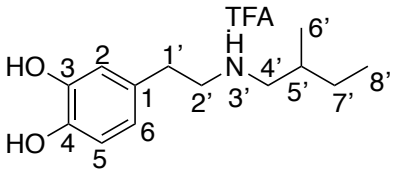
_N_-(3,4-Bis(benzyloxy)phenethyl)-2-methylbutan-1-amine (108)[30]



Under anhydrous conditions, a solution of *N*-(3,4-bis(benzyloxy)phenethyl)-2-methylbutanamide (65 mg, 0.16 mmol) in THF (4.5 mL) was prepared. Boron trifluoride diethyl etherate (9.9 μL, 0.080 mmol) was added and the reaction heated under reflux for 10 min. Borane dimethyl sulfide complex (2.0 M in THF, 240 μL, 0.48 mmol) was added and the reaction refluxed for a further 3 h. The reaction mixture was cooled to 0

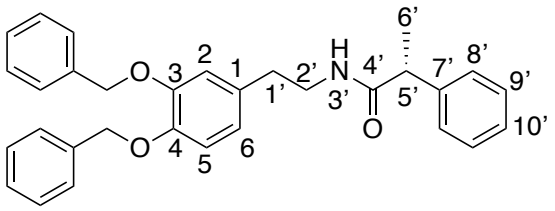
°C, 10% HCl (2.2 mL) was added and the reaction mixture stirred for 1 h. The reaction mixture was warmed to room temperature and stirred for a further 18 h. The pH of the reaction mixture was adjusted to 13 and the reaction extracted into dichloromethane (3 x 10 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under reduced pressure to a purple oil. This crude product was taken through to the next reaction step without further purification. *m/z* [HRMS ES+] found $[M+H]^+$ 404.2586; $C_{27}H_{34}NO_2$ requires 404.2589.

_N_-(3,4-Dihydroxyphenethyl)-2-methylbutan-1-amine.TFA (109)[30]



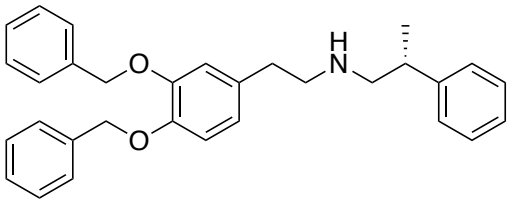
A solution of *N*-(3,4-bis(benzyloxy)phenethyl)-2-methylbutan-1-amine (40 mg, 0.10 mmol) in methanol (5 mL) and 37% HCl (1 mL) was prepared and heated under reflux for 24 h. Solvents were then removed under reduced pressure and the resulting residue purified by preparative HPLC (method 8) to give the desired product as a pale brown solid (TFA salt, 26 mg, 81%). ν_{max}/ cm^{-1} (thin film): 2976, 2860, 1666, 1600, 1521; ^{1}H-NMR (600 MHz; $CDCl_3$) δ 6.76 – 6.71 (1H, m, 5-H), 6.70 – 6.66 (1H, m, 2-H), 6.62 – 6.56 (1H, m, 6-H), 3.16 (2H, t, *J* = 8.2 Hz, 2'-H), 3.02 – 2.93 (1H, m, 4'-H), 2.87-2.78 (3H, m, 4'-H and 1'-H), 1.79 – 1.74 (1H, m, 5'-H), 1.50 – 1.47 (1H, m, 7'-H), 1.28 – 1.22 (1H, m, 7'-H), 1.02 (3H, d, *J* = 6.7 Hz, 6'-H), 0.96 (3H, t, *J* = 7.5 Hz, 8'-H); ^{13}C-NMR (151 MHz, $CDCl_3$) δ 146.8 (C-3), 145.6 (C-4), 129.0 (C-1), 120.9 (C-6), 116.8 (C-5), 116.7 (C-2), 54.7 (C-4'), 50.9 (C-2'), 33.7 (C-1'), 32.6 (C-5'), 27.9 (C-7'), 17.0 (C-6'), 11.2 (C-8'); *m/z* [ES+] 223 ($[M+H]^+$, 100%); *m/z* [HRMS ES+] found $[M+H]^+$ 223.1567; $C_{13}H_{21}NO_2$ requires 223.1567.

(*R*)-*N*-(3,4-Bis(benzyloxy)phenethyl)-2-phenylpropanamide[30,148]



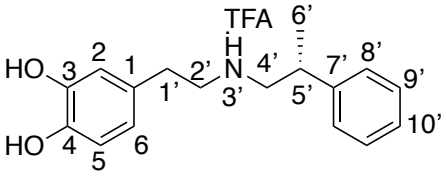
A solution of 3,4-di(benzyloxy)phenethylamine (400 mg, 0.92 mmol) and triethylamine (416 μL, 3.0 mmol) in dichloromethane (15 mL) was prepared under anhydrous conditions and stirred. (*R*)-2-Phenylpropionic acid (150 mg, 1.0 mmol) and thionyl chloride (73 μL, 1.0 mmol) were added and the reaction mixture stirred at room temperature for 30 min. Solvents were removed under reduced pressure and the resulting residue was re-suspended in dichloromethane (10 mL) and 1 M HCl (10 mL). The organic phase was washed with 1 M HCl (2 x 10 mL) and 1 M NaOH (3 x 10 mL), dried with Na_2SO_4 and concentrated under reduced pressure. The resulting residue was purified by silica chromatography (20-50% ethyl acetate in petroleum ether 40-60) to yield the product (178 mg, 41%) as a colourless oil. $[\alpha]_D^{26}$ 0.018 (c = 0.46, $CHCl_3$); ν_{max}/ cm^{-1} (thin film): 2950, 3050, 3000, 2970, 2910, 2845, 1660, 1610, 1550, 1500; ^{1}H-NMR (400 MHz; $CDCl_3$) δ 7.47 - 7.16 (15H, m, Ar*H*), 6.78 (1H, d, *J* = 8.1 Hz, 5-H), 6.70 – 6.67 (1H, m, 2-H), 6.49 – 6.46 (1H, m, 6-H), 5.13 (2H, s, PhCH_2O), 5.08 (2H, s, PhCH_2O), 3.50 - 3.38 (2H, m, 2'-H), 3.38 - 3.28 (1H, m, 5'-H), 2.63 - 2.56 (2H, m, 1'-H), 1.48 (3H, d, *J* = 7.2 Hz, 6'-H); ^{13}C-NMR (176 MHz, $CDCl_3$) δ 174.3 (C-4'), 149.1 (C-3 or C-4), 147.7 (C-3 or C-4), 141.3, 137.6, 137.4, 132.3, 129.0, 128.8, 128.6, 128.6, 128.0, 179.9, 127.8, 127.7, 127.5, 127.4, 71.6 (O*C*H_2Ph), 71.4 (O*C*H_2Ph), 47.2 (C-5'), 40.8 (C-2'), 35.1 (C-1'), 18.5 (C-6'); *m/z* [ES+] 467 ([M+H]$^+$, 100%); *m/z* [HRMS ESI+] found [M+H]$^+$ 466.2380; $C_{31}H_{32}NO_2$ requires 466.2377.

(*R*)-*N*-(3,4-Bis(benzyloxy)phenethyl)-2-phenylpropan-1-amine[30,148]



Under anhydrous conditions, a solution of (*R*)-*N*-(3,4-bis(benzyloxy)phenethyl)-2-phenylpropanamide (160 mg, 0.34 mmol) in THF (10 mL) was prepared. Boron trifluoride etherate (21 μL, 0.17 mmol) was added and the reaction mixture stirred at under reflux for 10 min. Boron dimethyl sulphide complex (2M in THF, 467 μL, 1.2 mmol) was added and the reaction mixture stirred under reflux for 3 h. The solution was then cooled to 0 °C and 10% HCl (8 mL) was added, and the reaction mixture stirred at 0 °C for 1 h then at rt for 2 h. The pH of the solution was adjusted to 13 by addition of 2 M NaOH and the product extracted into CH_2Cl_2 (3 x 10 mL). The organic phase was dried (Na_2SO_4) and concentrated to give a pale-yellow oil (157 mg, 100%). The product was taken directly through to the next stage without further purification.

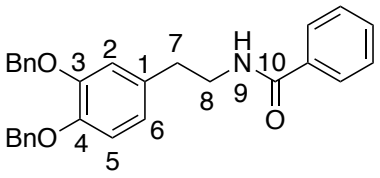
(*R*)-4-(2-((2-Phenylpropyl)amino)ethyl)benzene-1,2-diol.TFA (110)[30,148]



A solution of (*R*)-*N*-(3,4-bis(benzyloxy)phenethyl)-2-phenylpropan-1-amine (157 mg, 0.34 mmol) in MeOH (15 mL) and 37% HCl (3 mL) was prepared under anhydrous conditions. The reaction mixture was heated under reflux for 18 h then solvents were removed under reduced pressure. The resulting residue was purified by preparative, HPLC (method 8) to isolate the product as a white solid (TFA salt, 13.6 mg, 15%, >99% e.e. as determined by chiral HPLC method 4). $[\alpha]_D^{25}$ 0.005 (c = 0.1, MeOH); υ_{max}/ cm^{-1} (thin film): 3000, 3020, 2860, 1660, 1590, 1500; ^{1}H-NMR (600 MHz; D_2O) δ 7.38 – 7.34 (2H, m, 8'-H or 9'-H), 7.33 – 7.29 (1H, m, 10'-H), 7.28 – 7.25 (2H, m, 8'-H or 9'-H), 6.81

– 6.77 (1H, m, 5-H), 6.71 – 6.68 (1H, m, 2-H), 6.61 – 6.57 (1H, m, 6-H), 3.27 - 3.19 (3H, m, 2'-H and 4'H), 3.17 - 3.06 (2H, m, 1'-H), 2.82 - 2.73 (2H, m, 4'-H and 5'-H), 1.25 (3H, d, J = 7.0 Hz, 6'-H); ^{13}C-NMR (151 MHz; D_2O) δ 144.8 (C-7'), 143.7 (C-3 or C-4), 141.9 (C-3 or C-4), 129.9, 129.2, 128.4, 127.8, 121.6 (C-6), 117.1 (C-2 or C-5), 116.9 (C-2 or C-5), 53.7 (C-4'), 49.1 (C-2'), 37.4 (C-5'), 31.0 (C-1'), 19.5 (C-6'); *m/z* [ES+] 273 ([M+H]$^+$, 100%); *m/z* [HRMS ESI+] found [M+H]$^+$ 272.1644; $C_{17}H_{22}NO_2$ requires 272.1645.

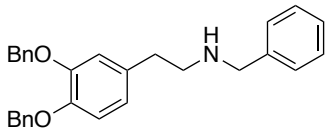
_N_-(3,4-Bis(benzyloxy)phenethyl)benzamide[30,149]



A solution of 2-(3,4-bis(benzyloxy)phenyl)ethan-1-amine (317 mg, 0.950 mmol), benzoic acid (128 mg, 1.04 mmol) and HATU (362 mg, 0.950 mmol) in DMF (5 mL) was prepared in anhydrous conditions and cooled to 0 °C. *N,N*-Diisopropylethylamine (332 μL, 1.90 mmol) was added dropwise and the reaction mixture allowed to room temperature. The reaction mixture was stirred for 18 h, then solvents removed under reduced pressure. The resulting residue was resuspended in ethyl acetate (15 mL), washed with 1 M HCl (2 x 15 mL), $NaHCO_3$ (2 x 15 mL) and brine (2 x 15 mL), dried and concentrated under reduced pressure. The resulting residue was purified by column chromatography (5 – 25% ethyl acetate in petroleum ether 40-60), to give the pure product as a white solid (290 mg, 69%). ν_{max}/ cm^{-1} (thin film): 3309, 3060, 3028, 2927, 2858, 1637, 1601, 1576.5, 1534, 1510; ^{1}H-NMR (700 MHz; $CDCl_3$) δ 7.68 - 7.64 (2H, m, COC(C*H*)$_2$), 7.50 - 7.26 (13H, m, Ar*H*), 6.92 – 6.88 (1H, m, 5-H), 6.84 – 6.81 (1H, m, 2-H), 6.76 – 6.73 (1H, m, 6-H), 6.10 – 6.05 (1H, m, N*H*), 5.14 (2H, s, PhC*H*$_2$O), 5.11 (2H, s, PhC*H*$_2$O), 3.65 (2H, q, J = 5.9 Hz, 6.8 Hz, C*H*$_2$NH), 2.83 (2H, t, J = 6.8 Hz, C*H*$_2$CHNH); ^{13}C-NMR (176 MHz; $CDCl_3$) δ 167.5 (C-10), 149.3 (C-3 or C-4), 147.9 (C-3 or C-4), 137.5, 137.3, 134.8, 132.4, 131.6, 128.7, 128.6, 128.0, 127.9, 127.5, 127.5,

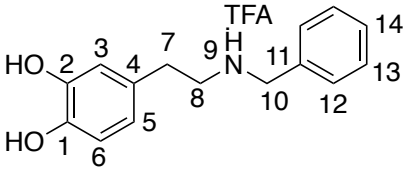
126.9, 121.8 (C-6), 115.9 (C-2 or C-5), 115.7 (C-2 or C-5), 71.6 (O*C*H_2Ph), 71.5 (O*C*H_2Ph), 41.2 (C-8), 35.2 (C-7); *m/z* [ES+] 438 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 438.2050; $C_{29}H_{27}NO_2$ requires 438.2069.

_N_-Benzyl-2-(3,4-bis(benzyloxy)phenyl)ethan-1-amine[30,149]



A solution of *N*-(3,4-bis(benzyloxy)phenethyl) benzamide (250 mg, 0.57 mmol) in THF (20 mL) was prepared under anhydrous conditions. Boron trifluoride diethyl etherate (36 μL, 0.29 mmol) was added and the reaction stirred under reflux for 10 min. Boron dimethyl sulfide complex (2M in THF, 860 μL, 1.70 mmol) was added and the reaction mixture stirred for a further 3 h. The mixture was then cooled to 0 °C, 1 M HCl (10 mL) was added and the reaction stirred for 1 h. The reaction was then warmed to room temperature and stirred overnight. The pH was adjusted to 13 by addition of 2 M NaOH, the solution extracted with dichloromethane (3 x 20 mL). The organic phase was then dried with Na_2SO_4 and concentrated under reduced pressure to give the crude product as a grey oil which was taken through to the next step without further purification.

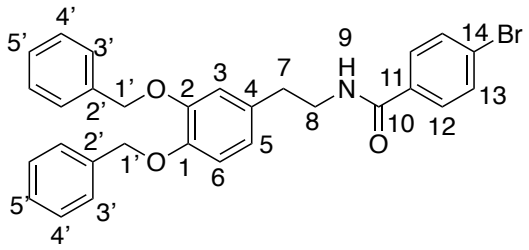
4-(2-(Benzylamino)ethyl)benzene-1,2-diol.TFA (190)[30,149]



A solution of *N*-benzyl-2-(3,4-bis(benzyloxy)phenyl)ethan-1-amine (240 mg, 0.57 mmol) in MeOH (30 mL) and 37% HCl (6 mL) was heated under reflux for 18 h. Solvents were removed under reduced pressure and the resulting residue was purified by preparative HPLC (method 9), lyophilised and the pure product isolated as a white solid (TFA salt, 57 mg, 29%). ν_{max}/ cm^{-1} (thin film): 3039, 1661, 1528; ^{1}H-NMR (700 MHz; D_2O) δ 7.47 - 7.37 (5H, m, 12-H,13-H,14-H), 6.84 (1H, d, *J* = 8.1 Hz, 6-H), 6.76 (1H, d,

J = 2.2 Hz, 3-H), 6.68 (1H, dd, *J* = 8.1, 2.2 Hz, 5-H), 4.19 (2H, s, 10-H), 3.23 (2H, t, *J* = 7.6 Hz, 8-H), 2.86 (2H, t, *J* = 7.6 Hz, 7-H); ^{13}C-NMR (176 MHz; D_2O) δ 144.8 (C-1 or C-2), 143.6 (C-1 or C-2), 131.1 (C-11), 130.3, 130.2, 129.8, 129.6, 121.7, 117.1 (C-3 or C-6), 117.0 (C-3 or C-6), 50.8 (C-10), 48.5 (C-8), 31.4 (C-7); *m/z* [ES+] 244 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 244.1332; $C_{15}H_{17}NO_2$ requires 244.1338.[149]

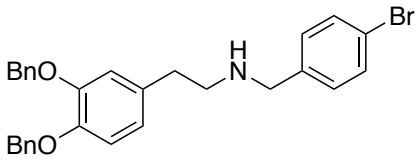
_N_-(3,4-Bis(benzyloxy)phenethyl)-4-bromobenzamide[30]



A solution of 2-(3,4-bis(benzyloxy)phenyl)ethan-1-amine (200 mg, 0.46 mmol), 4-bromobenzoic acid (110 mg, 0.55 mmol) and HATU (175 mg, 0.46 mmol) in DMF (5 mL) was prepared under anhydrous conditions and cooled to 0 °C. *N,N*-Diisopropylethylamine (160 μL, 0.92 mmol) was added dropwise and the reaction mixture allowed to room temperature. The reaction mixture was stirred for 18 h, then solvents removed under reduced pressure. The resulting residue was resuspended in ethyl acetate (15 mL), washed with 1 M HCl (2 x 15 mL), $NaHCO_3$ (2 x 15 mL) and brine (2 x 15 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under reduced pressure. The resulting residue was purified by column chromatography (5 – 25% ethyl acetate in petroleum ether 40-60) to give the pure product as a white solid (120 mg, 61%). ν_{max}/ cm^{-1} (thin film): 3287, 3061, 3029, 2916, 2848, 1635, 1603, 1587, 1540, 1513; ^{1}H-NMR (700 MHz; $CDCl_3$) δ 7.55 – 7.27 (14H, m, Ar*H*), 6.93 – 6.90 (1H, m, 6-H), 6.89 (1H, d, *J* = 8.5 Hz, 3-H), 6.72 (1H, dd, *J* = 8.1, 2.0 Hz, 5-H), 5.98 (0.8H, m, NH), 5.14 (2H, s, 1'-H), 5.11 (2H, s, 1'-H), 3.63 (2H, q, *J* = 6.4 Hz, 8-H), 2.82 (2H, t, *J* = 6.4 Hz, 7-H); ^{13}C-NMR (176 MHz; $CDCl_3$) δ 166.5 (C-10), 149.3 (C-1 or C-2), 147.9 (C-1 or C-2), 137.4 (C-2'), 137.3 (C-2'), 133.6, 132.2, 131.9, 128.6, 128.6, 128.0, 127.9, 127.5, 127.5, 126.2, 121.8 (C-5), 115.9 (C-3), 115.6 (C-6), 71.6 (C-1'),

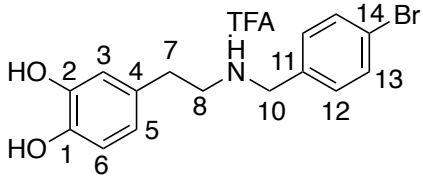
71.5 (C-1'), 41.2 (C-8), 35.1 (C-7); *m/z* [ES$^+$] 516 ([M+H]$^+$, 100%); *m/z* [HRMS ESI$^+$] found [M+H]$^+$ 516.1171; $C_{29}H_{26}BrNO_2$ requires 516.1174.

2-(3,4-Bis(benzyloxy)phenyl)-*N*-(4-bromobenzyl)ethan-1-amine[30]



A solution of *N*-(3,4-bis(benzyloxy)phenethyl)-4-bromobenzamide (120 mg, 0.23 mmol) in THF (10 mL) was prepared under anhydrous conditions. Boron trifluoride diethyl etherate (16 μL, 0.13 mmol) was added and the reaction stirred under reflux for 10 min. Boron dimethyl sulfide complex (2.0 M in THF, 350 μL, 0.69 mmol) was added and the reaction mixture stirred for a further 3 h. The mixture was then cooled to 0 °C, 1 M HCl (5 mL) was added and the reaction stirred for 1 h. The reaction was then warmed to room temperature and stirred overnight. The pH was adjusted to 13 by addition of 2 M NaOH, the solution extracted with dichloromethane (3 x 20 mL). The organic phase was then dried with Na_2SO_4 and concentrated under reduced pressure to give the crude product as a colourless oil which was taken through to the next step without further purification.

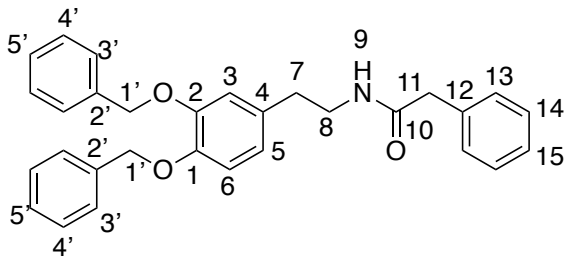
4-(2-((4-Bromobenzyl)amino)ethyl)benzene-1,2-diol.TFA (191)[30]



A solution of 2-(3,4-bis(benzyloxy)phenyl)-*N*-(4-bromobenzyl)ethan-1-amine (120 mg, 0.24 mmol) in MeOH (30 mL) and 37% HCl (6 mL) was heated under reflux for 18 h. Solvents were removed under reduced pressure and the resulting residue was purified by preparative HPLC (method 9), lyophilised and the pure product isolated as a white

solid as the TFA salt (21 mg, 27%). ν_{max}/ cm^{-1} (thin film): 3110, 2823, 1664, 1604, 1579, 1524; ^{1}H-NMR (700 MHz; $CDCl_3$) δ 7.62 - 7.59 (2H, m, 13-H), 7.31 - 7.27 (2H, m, 12-H), 6.86 – 6.83 (1H, m, 6-H), 6.76 – 6.72 (1H, m, 3-H), 6.71 – 6.66 (1H, m, 5-H), 4.16 (2H, s, 10-H), 3.23 (2H, t, *J* = 7.6 Hz, 8-H), 2.86 (2H, t, *J* = 7.6 Hz, 7-H); ^{13}C-NMR (176 MHz; $CDCl_3$) δ 144.8 (C-1 or C-2), 143.6 (C-1 or C-2), 132.8 (Ar*C*), 132.1 (Ar*C*), 130.2 (Ar*C*), 129.5 (Ar*C*), 123.9 (C-5 or C-14), 121.7 (C-5 or C-14), 117.1 (C-3 or C-6), 117.0 (C-3 or C-6), 50.8 (C-10), 48.5 (C-8), 31.4 (C-7); *m/z* [ES$^+$] 322 ([M+H]$^+$, 100%); *m/z* [HRMS ESI$^+$] found [M+H]$^+$ 332.0435; $C_{15}H_{16}BrNO_2$ requires 322.0432.

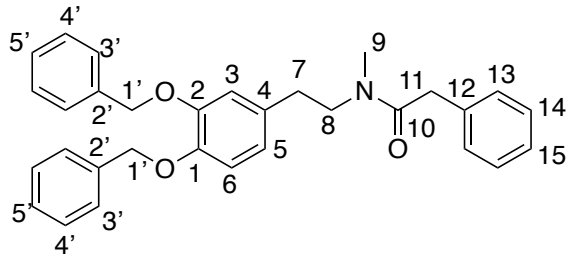
_N_-(3,4-Bis(benzyloxy)phenethyl)-2-phenylacetamide (211)[30]



A solution of 3,4-di(benzyloxy)phenethylamine (360 mg, 0.83 mmol) and triethylamine (330 μL, 2.4 mmol) in dichloromethane (12 mL) was prepared under anhydrous conditions and stirred. Phenylacetic acid (100 μL, 0.80 mmol) was added, then thionyl chloride (73 μL, 1.0 mmol) and the reaction mixture stirred for 20 min at rt. Solvents were removed under reduced pressure and the resulting residue resuspended in dichloromethane (12 mL). The organic phase was washed (3 x 1 M HCl then 3 x 1 M NaOH) and then the organic phases dried with Na_2SO_4 and concentrated under reduced pressure. The resulting residue was purified by column chromatography (20 – 50% ethyl acetate in petroleum ether 40-60) to give the product as a white solid (100 mg, 27%). ν_{max}/ cm^{-1} (thin film): 3297, 3062, 3029, 2934, 2869, 1720, 1646, 1603, 1585, 1511; ^{1}H-NMR (600 MHz; $CDCl_3$) δ 7.47 – 7.42 (4H, m, Ar*H*), 7.40 – 7.22 (9H, m, Ar*H*), 7.15 – 7.12 (2H, m, Ar*H*), 6.78 (1H, d, *J* = 8.1 Hz, 6-H), 6.68 (1H, d, *J* = 2.0 Hz, 3-H), 6.50 (1H, dd, *J* = 8.1 Hz, 2.0 Hz), 5.34 – 5.28 (1H, m, 9-H), 5.14 (2H, s, 1'-H), 5.10 (2H, s, 1'-H), 3.51 (2H, s, 11-H), 3.39 (2H, q, *J* = 7.0 Hz, 8-H), 2.62 (2H, t, *J* = 7.0 Hz, 7-H);

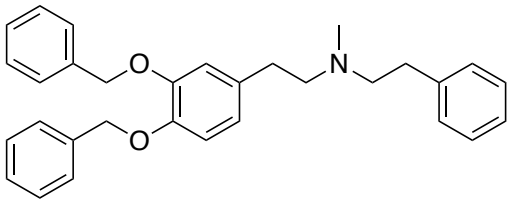
^{13}C-NMR (151 MHz, $CDCl_3$) δ 171.1 (C-10), 149.1 (C-2), 147.7 (C-1), 137.5, 137.3, 134.8, 133.9, 132.1, 129.6, 129.5, 129.1, 128.7, 128.6, 128.6, 128.0, 127.9, 127.5, 127.4, 127.4, 127.3, 121.6 (C-5), 115.6 (C-3), 115.4 (C-6), 71.5 (C-1'), 71.3 (C-1'), 43.9 (C-11), 40.8 (C-7), 35.0 (C-8); *m/z* [ES$^+$] 452 ([M+H]$^+$, 100%); *m/z* [HRMS ES+] found [M+H]$^+$ 452.2222; $C_{30}H_{30}NO_3$ requires 452.2220.

_N_-(3,4-Bis(benzyloxy)phenethyl)-_N_-methyl-2-phenylacetamide (212)[30]



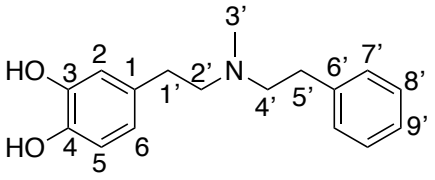
A solution of *N*-(3,4-bis(benzyloxy)phenethyl)-2-phenylacetamide (97 mg, 0.21 mmol) in DMF (2 mL) was prepared under anhydrous conditions. The solution was cooled to 0 °C and stirred, then sodium hydride (60% in mineral oil, 13.5 mg, 0.32 mmol) was added. Iodomethane (26 μL, 0.42 mmol) was added and the reaction mixture allowed to warm to rt and stirred for 18 h. The reaction was quenched by addition of water (500 μL) and solvents were removed under reduced pressure. The residue was resuspended in water (10 mL) and extracted with diethyl ether (3 x 10 mL). The organic phases were combined, washed with saturated aqueous $NaHCO_3$ (2 x 10 mL) and brine (2 x 10 mL). The organic phase was then dried with Na_2SO_4 and concentrated under reduced pressure. The resulting residue was purified by column chromatography (20% ethyl acetate in petroleum ether 40-60) to give the desired product as a white solid (51 mg, 52%). ν_{max}/ cm^{-1}(thin film): 3062, 3029, 2925, 2856, 1641, 1604, 1512; ^{1}H-NMR (600 MHz; $CDCl_3$) δ 7.37 – 7.33 (4H, m, Ar*H*), 7.28 – 7.22 (4H, m, Ar*H*), 7.23 – 7.18 (4H, m, Ar*H*), 7.15 – 7.11 (2H, m, Ar*H*), 7.08 – 7.04 (1H, m, Ar*H*), 6.80 (1H, d, *J* = 8.3 Hz, 6-H), 6.73 (1H, d, *J* = 1.9 Hz, 3-H), 6.56 (1H, dd, *J* = 8.3, 1.9 Hz, 5-H), 5.05 (2H, s, 1'-H), 5.03 (2H, s, 1'-H), 3.56 (2H, s, 11-H), 3.47 (2H, t, *J* = 7.2 Hz, 8-H), 2.85 (3H, s, 9-H), 2.67 (2H, t, *J* = 7.2 Hz, 7-H); *m/z* [ES$^+$] 467 ([M+H]$^+$, 100%).

_N_-(3,4-Bis(benzyloxy)phenethyl)-_N_-methyl-2-phenylethan-1-amine (213)[30]



Under an inert atmosphere, a solution of *N*-(3,4-bis(benzyloxy)phenethyl)-*N*-methyl-2-phenylacetamide (45 mg, 0.10 mmol) in THF (4 mL) was prepared. Boron trifluoride diethyl etherate (7.0 μL, 0.050 mmol) was added, the reaction stirred and heated under reflux for 10 min. Borane dimethyl sulfide complex (2.0 M in THF, 160 μL, 0.30 mmol) was added dropwise and the reaction heated under reflux for 3 h. The reaction mixture was then cooled to 0 °C, quenched by addition of 10% HCl (2.5 mL) then allowed to warm to rt and stirred overnight. The pH of the solution was adjusted to 13 by addition of 6 M NaOH. This aqueous solution was extracted with dichloromethane (3 x 10 mL), then the organic phases combined, dried with Na_2SO_4 and concentrated under reduced pressure. The resulting residue, the crude product, was taken through to the next step without further purification.

4-(2-(Methyl(phenethyl)amino)ethyl)benzene-1,2-diol (210)[30]

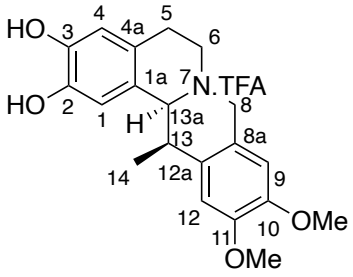


A solution of *N*-(3,4-bis(benzyloxy)phenethyl)-*N*-methyl-2-phenylethan-1-amine (40 mg, 0.089 mmol) in methanol (8 mL) and 37% HCl (2 mL) was prepared, stirred and heated under reflux for 24 h. Solvents were then removed under reduced pressure and the residue purified by preparative HPLC (method 11), then prepared by workup method 3 to give the pure product as a white solid (2.1 mg, 5.7 μmol, 6.0%). ν_{max}/ cm^{-1} (thin film): 3030, 1670, 1600, 1528; ^{1}H-NMR (600 MHz; D_2O) δ 7.38 - 7.34 (2H, m, 7'-H

or 8'-H), 7.32 - 7.28 (1H, m, 9'-H), 7.26 - 7.22 (2H, m, 7'-H or 8'-H), 6.85 – 6.82 (1H, m, 5-H), 6.80 – 6.77 (1H, m, 2-H), 6.70 – 6.66 (1H, m, 6-H), 3.49 - 3.42 (2H, m, 2'-H or 4'-H), 3.37 - 3.28 (2H, m, 2'-H or 4'-H), 3.07 - 2.92 (4H, m, 1'-H and 5'-H), 2.90 (3H, s, 3'-H); ^{13}C-NMR (176 MHz; D_2O) δ 144.9 (C-3), 143.7 (C-4), 136.5 (C-6), 129.7 (C-1), 129.3 (C-7' or C-8'), 129.1 (C-7' or C-8'), 128.0 (C-9'), 121.6 (C-6), 117.2 (C-2 or C-5), 117.0 (C-2 or C-5), 57.4 (C-2' or C-4'), 57.3 (C-2' or C-4'), 40.9 (C-3'), 30.3 (C-1' or C-5'), 29.7 (C-1' or C-5'); *m/z* [ES$^+$] 272 ([M+H]$^+$, 100%); *m/z* [HRMS ESI$^+$] found [M+H]$^+$ 272.1641, $C_{17}H_{22}NO_2$ requires 272.1645.

8.8.5 Further functionalization of enzymatic products with chemical transformations

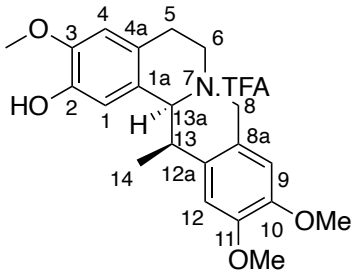
(13*R*,13a*S*)-9,10-Dimethoxy-13-methyl-5,8,13,13a-tetrahydro-6*H*-isoquinolino[3,2-*a*]isoquinoline-2,3-diol.TFA (129)[198]



An NCS-mediated reaction between dopamine HCl and 2-(3,4-dimethoxyphenyl)propanal was performed on a 5 mL, 10 mM scale, as detailed for the synthesis of (*S*)-1-((*R*)-1-(3,4-dimethoxyphenyl)ethyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol (**119**) then workup method 5 was performed. The residue was resuspended in formic acid (11.6 mL) and formaldehyde (37% in water, 8.4 mL). The reaction mixture was stirred at 50 °C for 2 h. Solvents were removed under reduced pressure and the residue purified by preparative HPLC (method 11) to give the product (TFA salt) as a yellow solid (5.3 mg, 24%). $[\alpha]_D^2$ -35.3 (*c* 0.83, MeOH)**;** ν_{max}/ cm^{-1} (thin film): 2980, 1672, 1614, 1521; ^{1}H-NMR (700 MHz; CD_3OD) δ 6.94 (1H, s, 12-H), 6.79 (1H, s, 9-H), 6.76

(1H, s, 1-H), 6.65 (1H, s, 4-H), 4.91 (1H, d, *J* = 4.0 Hz, 13a-H), 4.54 – 4.51 (2H, m, 8-H), 3.85 (3H, s, OC*H*$_3$), 3.83 (3H, s, OC*H*$_3$), 3.79 – 3.74 (2H, m, 13-H and 6-H), 3.51 (1H, dt, *J* = 3.8 Hz, 6-H), 3.19 – 3.12 (1H, m, 5-H), 2.90 (1H, dd, *J* = 17.2 Hz, 2.7 Hz, 5-H), 1.10 (3H, d, *J* = 7.3 Hz, 14-H); ^{13}C-NMR (176 MHz, CD_3OD) δ 150.9 (C-11), 150.1 (C-10), 146.7 (C-3), 146.7 (C-2), 131.1 (C-12a), 124.2 (C-4a), 121.7 (C-1a), 119.6 (C-8a), 115.9 (C-4), 113.0 (C-12), 112.8 (C-1), 109.7 (C-9), 65.6 (C-13a), 57.3 (C-8), 56.4 (O*C*H$_3$), 56.4 (O*C*H$_3$), 52.8 (C-6), 36.7 (C-13), 26.4 (C-5), 17.3 (C-14); *m/z* [ES+] 342 ([M+H]$^+$, 100%); *m/z* [HRMS ESI$^+$] found [M+H]$^+$ 342.1701; $C_{20}H_{25}NO_4$ requires 342.1670; retention time (achiral analytical HPLC method 2) = 8.9 min.

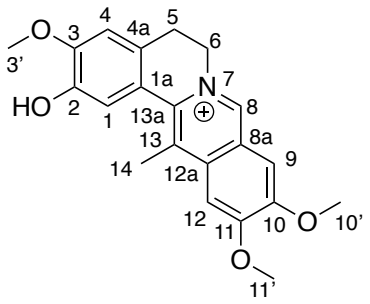
(13*R*,13a*S*)-3,10,11-Trimethoxy-13-methyl-5,8,13,13a-tetrahydro-6*H*-isoquinolino[3,2-*a*]isoquinolin-2-ol.TFA (130)[198]



A crude, lyophilised sample of (*S*)-1-((*R*)-1-(3,4-dimethoxyphenyl)ethyl)-6-methoxy-1,2,3,4-tetrahydroisoquinolin-7-ol was resuspended in formic acid (11.6 mL) and formaldehyde (37% in water, 8.4 mL). The reaction mixture was stirred at 50 °C for 2 h.[271] Solvents were removed under reduced pressure and the residue purified by preparative HPLC (method 11) to give the product as the TFA salt as a pale yellow solid (5.1 mg, 23%). $[\alpha]_D^{20}$ +0.92 (*c* 0.72, MeOH)**;** ν_{max}/ cm^{-1} (thin film): 3293, 2967, 2842, 1757, 1668, 1613; ^{1}H-NMR (700 MHz; CD_3OD) δ 6.95 (1H, s, 1-H), 6.83 – 6.81 (2H, m, 4-H and 12-H), 6.80 (1H, s, 9-H), 4.96 (1H, d, *J* = 4.1 Hz, 13a-H), 4.54 (2H, s, 8-H), 3.88 (3H, s, OC*H*$_3$), 3.87 (3H, s, OC*H*$_3$), 3.83 (3H, s, OC*H*$_3$), 3.82 – 3.77 (2H, m, 6-H and 13-H), 3.57 – 3.51 (1H, s, 6-H), 3.25 – 3.17 (1H, m, 5-H), 3.03 – 2.97 (1H, m, 5-H), 1.09 (3H, d, *J* = 7.4 Hz, 14-H); ^{13}C-NMR (176 MHz; CD_3OD) δ 150.9 (C-10 or C-11), 150.1 (C-3), 149.1 (C-10 or C-11), 147.7 (C-2), 131.1 (C-8a), 124.2 (C-4a), 123.0 (C-

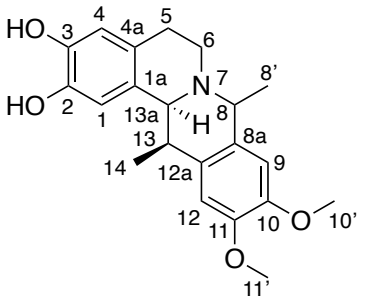
1a), 119.6 (C-12a), 113.0 (C-1), 112.8 (C-12), 112.3 (C-4), 109.7 (C-9), 65.5 (C-13a), 57.3 (C-8), 56.4 (O*C*H$_3$), 56.4 (O*C*H$_3$), 56.3 (O*C*H$_3$), 52.7 (C-6), 36.5 (C-13), 26.5 (C-5), 17.2 (C-14); *m/z* [ES$^+$] 356 ([M+H]$^+$, 100%); *m/z* [HRMS ESI$^+$] found [M+H]$^+$ 356.1854, $C_{21}H_{26}NO_4$ requires 352.1856; retention time (achiral analytical HPLC method 2) = 9.2 min.

2-Hydroxy-3,10,11-Trimethoxy-13-methyl-5,6-dihydroisoquinolino[3,2-*a*]isoquinolin-7-ium (131)



Under anhydrous conditions, a crude, lyophilised sample of (*S*)-1-((*R*)-1-(3,4-dimethoxyphenyl)ethyl)-6-methoxy-1,2,3,4-tetrahydroisoquinolin-7-ol (**125**) was resuspended in DMF (10 mL). Paraformaldehyde (6.0 mg, 0.20 mmol) was added. The reaction mixture was stirred at 120 ^oC for 18 h. Solvents were removed under reduced pressure and the resulting residue was purified by preparative HPLC (method 11) to give the product as a brown oil (1.7 mg, 4.8%). ν_{max}/ cm^{-1} (thin film): 3523, 3321, 3012, 1674, 1605, 1574, 1504; ^{1}H-NMR (700 MHz; CD_3OD) δ 9.32 (1H, s, 8-H), 7.62 (1H, s, 9-H), 7.55 (1H, s, 12-H), 7.33 (1H, s, 1-H), 7.08 (1H, s, 4-H), 4.68 (2H, t, *J* = 5.9 Hz, 6-H), 4.16 (3H, s, 11'-H), 4.06 (3H, s, 10'-H), 3.98 (3H, s, 3'-H), 3.14 (2H, t, *J* = 5.9 Hz, 5-H), 2.97 (3H, s, 14-H); ^{13}C-NMR (176 MHz; CD_3OD) δ 159.6 (C-11), 154.1 (C-10), 151.2 (C-3), 146.5 (C-2), 144.8 (C-8), 138.9 (C-12a or C-8a), 138.7 (C-12a or C-8a), 131.4 (C-4a), 129.3 (C-13), 123.1 (C-13a), 121.0 (C-1a), 118.5 (C-1), 111.7 (C-4), 107.9 (C-9), 104.2 (C-12), 57.9 (C-6), 57.3 (C-11'), 56.9 (C-10'), 56.5 (C-3'), 28.7 (C-5), 18.1 (C-14); *m/z* [ES$^+$] 352 ([M]$^+$, 100%); *m/z* [HRMS ESI$^+$] found [M]$^+$ 352.1543, $C_{21}H_{22}NO_4$ requires 352.1534.

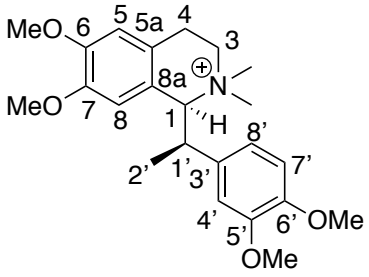
(8*SR*,13*R*,13a*S*)-10,11-Dimethoxy-8,13-dimethyl-5,8,13,13a-tetrahydro-6*H*-isoquinolino[3,2-*a*]isoquinoline-2,3-diol.TFA (140)[198]



A solution of (*S*)-1-((*R*)-1-(3,4-dimethoxyphenyl)ethyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol (16.5 mg, 0.05 mmol) in formic acid (2.1 mL), water (1.3 mL) and acetaldehyde (780 μL, 14.0 mmol) was prepared under anhydrous conditions and stirred at 50 °C for 24 h. Solvents were removed under reduced pressure and the resulting residue was resuspended in a minimum volume of methanol and purified by preparative HPLC (method 12) to give the product (TFA salt, not drawn) as brown solid (4.7 mg, 21%) as a mixture of two epimers (8*S*,13*R*,13a*S*):(8*R*,13*R*,13a*S*) in a ratio of 3:1. Stereochemistry at C-13 and C-13a was assigned based upon precedent of NCS reactions involving α-methyl substituted aldehydes and stereochemistry at C-8 was assigned based upon the presence or absence of NOE between 8-H and 13a-H. $[\alpha]_D^{20}$ -24.6 (*c* 0.45, MeOH)**;** ν_{max}/ cm^{-1} (thin film): 3449, 3011, 1673, 1650, 1516; ^{1}H-NMR (700 MHz; CD_3OD) δ 6.92 (0.3H, s, 12-H_{min}), 6.88 (1H, s, 12-H_{maj}), 6.85 (0.3H, s, 9-H_{min}), 6.82 – 6.79 (1.3H, m, 9-H_{maj} and C-1_{min}), 6.77 (1H, s, 1-H_{maj}), 6.67 (0.3H, s, 4-H_{min}), 6.65 (1H, s, 4-H_{maj}), 5.12 (1H, d, *J* = 4.1 Hz, 13a-H_{maj}), 4.18 (0.3 H, dd, *J* = 12.1, 5.0 Hz, 6-*H*H_{min}), 3.87 (0.9H, s, 11'-H or 10'-H), 3.86 – 3.83 (7H, m, 11'-H and 10'-H), 3.79 – 3.77 (0.3H, m, 6-H*H*$_{min}$), 3.75 – 3.71 (0.3 H, m, 13-H_{min}), 3.69 – 3.64 (1H, m, 13-H_{maj}), 3.60 – 3.55 (1H, m, 6-*H*H_{maj}), 3.52 (1H, td, *J* = 12.3, 2,8 Hz, 6-H*H*$_{maj}$), 3.44 – 3.37 (0.3H, m, 5-*H*H_{min}), 3.15 – 3.08 (1H, m, 5-*H*H_{maj}), 2.94 (0.3H, dd, *J* = 15.1, 3.3 Hz, 5-H*H*$_{min}$), 2.90 (1H, d, *J* = 17.1 Hz, 5-H*H*$_{maj}$), 1.85 (1H, d, *J* = 6.8 Hz, 8'-H_{min}), 1.69 (3H, d, *J* = 6.8 Hz, 8'-H_{maj}), 1.17 (1H, d, *J* = 7.3 Hz, 14-H_{min}), 1.04 (3H, d, *J* = 7.3 Hz, 14-H_{maj}). ^{13}C-NMR (171 MHz; CD_3OD) δ 150.8 (C-10_{maj} or C-11_{maj}), 150.5 (C-10_{min} or C-11_{min}),

150.1 (C-10_{min} or C-11_{min}), 150.1 (C-10_{maj} or C-11_{maj}), 146.7 (C-2 or C-3), 146.5 (C-2 or C-3), 131.2 (C-$12a_{min}$), 130.1 (C-$12a_{maj}$), 126.2 (C-$8a_{maj}$), 124.6 (C-$4a_{maj}$ and C-$8a_{min}$), 124.3 (C-$1a_{min}$ or C-$4a_{min}$), 122.3 (C-$1a_{maj}$), 122.0 (C-$1a_{min}$ or C-$4a_{min}$), 115.9 (C-4_{min}), 115.8 (C-4_{maj}), 113.3 (C-1_{maj}), 113.0 (C-1_{min}), 112.8 (C-12_{maj}), 112.4 (C-12_{min}), 110.3 (C-9_{maj}), 110.2 (C-9_{min}). 62.0 (C-8), 58.2 (C-$13a_{maj}$), 56.6 (C-10' or C-11'), 56.5 (C-10' or C-11'), 56.4 (C-10' or C-11'), 52.0 (C-6_{min}), 50.1 (C-6_{maj}), 49.4 (C-$13a_{min}$), 26.6 (C-5), 26.6 (C-5), 19.5 (C-$8'_{min}$), 17.3 (C-14_{maj}), 17.0 (C-14_{min}), 16.9 (C-$8'_{maj}$); *m/z* [ES+] 356 ($[M+H]^+$, 100%); *m/z* [HRMS ES+] found $[M+H]^+$ 356.1855; $C_{21}H_{26}NO_4$ requires 356.1856; retention time (achiral analytical HPLC method 2) = 8.6 min.

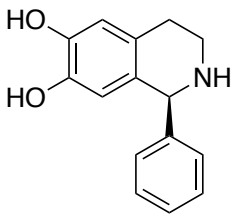
(*S*)-1-((*R*)-1-(3,4-Dimethoxyphenyl)ethyl)-6,7-dimethoxy-2,2-dimethyl-1,2,3,4-tetrahydroisoquinolin-2-ium (145)



An NCS-mediated reaction between dopamine HCl and 2-(3,4-dimethoxyphenyl)propanal was performed on a 5 mL, 10 mM scale as detailed for the synthesis of (*S*)-1-((*R*)-1-(3,4-dimethoxyphenyl)ethyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol, followed by lyophilisation. Under anhydrous conditions, the resulting residue was resuspended in methanol (5 mL) and stirred. Potassium carbonate (13.8 mg, 0.10 mmol) was added, followed by the dropwise addition of iodomethane (31 μL, 0.5 mmol). The reaction mixture was stirred at rt for 18 h. Solvents were removed under reduced pressure and the resulting residue purified by preparative HPLC (method 12) to give the product as a while solid (2.9 mg, 8.1%). $[\alpha]_D^{21}$ -11.3 (*c* 0.38, MeOH)**;** ν_{max}/ cm^{-1} (thin film): 3414, 2938, 2839, 1676, 1610, 1517; ^{1}H-NMR (700 MHz; CD_3OD) δ 6.94 (1H, s, 8-H), 6.83 – 6.80 (2H, m, C-5 and C-7'), 6.64 – 6.54 (1H, m, C-8'), 5.86 (1H, s, C-4'), 4.73 – 4.69 (1H, m, 1-H), 3.91 (3H, s, 7'-H), 3.89 – 3.86 (1H, m, 1'-H), 3.85 (3H, s, 6'-

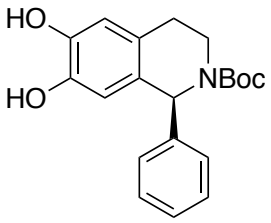
H), 3.76 (3H, s, 5'-OCH_3), 3.49 – 3.45 (6H, 2 x s, NCH_3 and 6'-OCH_3), 3.25 – 3.15 (2H, m, 3-H), 3.07 (3H, s, NCH_3), 3.01 – 2.93 (1H, m, 4-*H*H), 2.43 (1H, dd, *J* = 18.0, 6.5 Hz, 4-H*H*), 1.52 (3H, d, *J* = 7.3 Hz, 2'-H); ^{13}C-NMR (176 MHz; CD_3OD) δ 151.7 (C-6), 150.3 (C-6'), 149.9 (C-5'), 149.2 (C-7), 133.8 (C-8'), 124.3 (C-5a), 122.6 (C-8'), 120.8 (C-8a), 113.7 (C-8), 112.8 (C-5 or C-7'), 112.7 (C-5 or C-7'), 112.4 (C-4'), 77.1 (C-1), 56.7 (C-7'), 56.4 (C-6'), 56.2 (5'-O*C*H_3), 55.9 (6'-O*C*H_3 or N*C*H_3), 55.4 (C-3), 53.2 (6'-O*C*H_3 or N*C*H_3), 53.0 (N*C*H_3), 40.8 (C-1'), 24.1 (C-4 or C-1), 24.1 (C-4 or C-1); *m/z* [ES$^+$] 387 ([M]$^+$; *m/z* [HRMS ESI$^+$] found [M]$^+$ 386.2323, $C_{23}H_{32}NO_4$ requires 386.2323; retention time (achiral analytical HPLC method 2) = 10.9 min.

(1*S*)-1-Phenyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol (157)



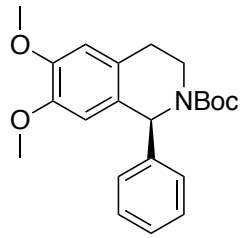
A solution of dopamine·HCl (37.9 mg, 0.2 mmol) and sodium ascorbate (39.6 mg, 0.2 mmol) in HEPES buffer (50 mM, pH 7.5, 18 mL) was prepared. To this, a solution of benzaldehyde (0.4 mmol, 40.8 μL) in MeCN (2 mL) was added and 4 mg of *Tf*NCS-M97V (10 mg mL^{-1} in 20 mM Tris, 50 mM NaCl, pH 7.5). The reaction mixture was stirred for 24 h at 37 °C. The reaction mixture was extracted into ethyl acetate (3 x 20 mL), dried and concentrated. The residue was resuspended in 1 M HCl (20 mL) and washed with dimethyl carbonate (1 x 5 mL). The aqueous layer was dried under reduced pressure. The product was resuspended in sat. aq. $NaHCO_3$, extracted into EtOAc (3 x 20 mL), dried (Na_2SO_4) and concentrated under reduced pressure to give the crude product as a yellow oil (30.0 mg, 62%) and the identity confirmed by ^{1}H-NMR analysis (matching that of the isolated product described on p282).

***tert*-Butyl-6,7-dihydroxy-1-phenyl-3,4-dihydroisoquinoline-2(1*H*)-carboxylate (181)**[30]



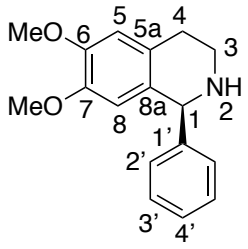
(1*S*)-1-Phenyl-1,2,3,4-tetrahydroisoquinoline-6,7-diol (30 mg, 0.12 mmol) was resuspended in THF (4 mL) and sat. aq. $NaHCO_3$ (4 mL). Boc anhydride (54 mg, 0.25 mmol) was added, and the reaction mixture stirred at rt (24 h). The reaction mixture was extracted into ethyl acetate (3 x 15 mL), dried (Na_2SO_4) and concentrated under reduced pressure. The product was taken through to the next step without further purification.

***tert*-Butyl-(*S*)-6,7-dimethoxy-1-phenyl-3,4-dihydroisoquinoline-2(1*H*)-carboxylate (189)**[229]



A solution of *tert*-butyl 6,7-dihydroxy-1-phenyl-3,4-dihydroisoquinoline-2(1*H*)-carboxylate (28 mg, 0.082 mmol) and base (potassium carbonate or caesium carbonate) (0.25 mmol) in methanol (5 mL) was prepared. Iodomethane (100 μL, 1.6 mmol) was added and the reaction mixture stirred for 24 h at room temperature. Solvents were removed under reduced pressure and the resulting residue azeotroped with methanol (3 x 10 mL). The product was partially purified by silica chromatography (100% dichloromethane) to give the product as a yellow oil (6.5 mg, 21%). The crude product was taken through to the next step without further purification.

(*S*)-6,7-Dimethoxy-1-phenyl-1,2,3,4-tetrahydroisoquinoline (143)[250]



A solution of *tert*-butyl (*S*)-6,7-dimethoxy-1-phenyl-3,4-dihydroisoquinoline-2(1*H*)-carboxylate (6.5 mg, 0.018 mmol) in dichloromethane (3.2 mL) and trifluoroacetic acid (0.8 mL) was prepared and stirred at room temperature for 5 min. Solvents were removed under reduced pressure, then resuspended in 50% MeOH in water and concentrated. The resulting residue was purified by semi-preparative HPLC or preparative HPLC (method 11) to give the desired product in 6% or 20% e.e. as determined by chiral HPLC methods 6 and 7, dependent on which base was used for the reaction (6% for potassium carbonate and 20% for cesium carbonate). ^{1}H-NMR (700 MHz; CD_3OD) δ 7.50 – 7.47 (3H, m, Ar*H*), 7.36 – 7.34 (2H, m, Ar*H*), 6.89 (1H, s, 5-H), 6.34 (1H, s, 8-H), 5.69 (1H, s, 1-H), 3.85 (3H, s, OCH_3), 3.59 (3H, s, OCH_3), 3.51 – 3.41 (2H, m, 3-H), 3.25 – 3.18 (1H, m, 4-*H*H), 3.14 – 3.08 (1H, m, 4-H*H*); ^{13}C-NMR (176 MHz; CD_3OD) δ 150.9 (C-6 or C-7), 149.8 (C-6 or C-7), 137.6 (C-1'), 131.0, 130.9, 130.3, 125.8, 124.0, 112.6 (C-5), 111.9 (C-8), 60.5 (C-1), 56.4 (OCH_3), 56.3 (OCH_3), 40.6 (C-3), 25.7 (C-4); *m/z* [ES+] 271 ([M+H]$^+$, 100%).

8.9 Crystallographic methods

8.9.1 Crystal preparation

Purified protein samples were concentrated to around 12 mg mL^{-1} in buffer (20 mM TRIS, 50 mM NaCl, pH 7.5. Initial crystallisation screening was performed using commercial 96-well plates (JCSG, PACT, Structure 1 + 2 and Index). For each well, two sitting drops were prepared using a Mosquito® robot (TTP Labtech) at 1:1 or 2:1 precipitant to protein.

Crystal optimisation was performed using either the hanging drop method or by preparing four corner screens. For the hanging drop method, two drops were prepared per well with 1:1 and 2:1 protein to precipitant, with drop sizes of 2 μL and 3 μL, respectively with 500 μL crystallisation solution in the well. Four corner screens involve preparing multi-well crystallisation plates, with minor changes in buffer, salt or precipitant concentrations around the promising crystallisation condition. Four corner screens were prepared as 96 well sitting drop plates or 24 well hanging drop plates.

8.9.2 Data collection and analysis

Diffraction data was collected from Diamond light source beamlines (i04 and i24). Data was analysed using CCP4i2 with the following software packages: XDS Dials[172] for image processing, Phaser[253] for molecular replacement, COOT[171] for model building and REFMAC5[174] for refinement with TLS restraints. Figures were prepared using UCSF Chimera[32] and CCP4mg.[272]

8.10 Computational methods

8.10.1 Data analysis

For data analysis of conversions of enzymatic assays, the errors given are standard deviations. Figures were prepared using Python with the matplotlib software package.

8.10.2 Energy minimization

Small molecule reaction intermediates, reaction products and substrates for docking were energy minimized using the Avogadro[273] software package using the universal force field (UFF).[274]

8.10.3 Molecular docking

AutoDockTools (Scripps) was used to define rotatable bonds in ligands and to prepare pdbqt files for docking. Docking was performed using AutoDock Vina (Scripps).[255]

Docking of wild-type *Tf*NCS was performed using PDB: 5NON, subunit A with ligands removed with box size (25, 25, 35) and box location (-17.95, -7.43, 16.19). Docking of *Tf*NCS-M97F was performed using an energy minimized structure of the variant (using the CHARMM GUI) based upon the structure of *Tf*NCS-M97V (PDB: 6Z82). A box size of (25, 25, 35) was used with a box centre of (-4, 14.4, 9.6).

8.10.4 Sequence alignment

Multiple sequence alignment of NCSs and homologs was performed using MAFFT[275] and pairwise sequence alignment was performed using EMBOSS Needle.[276] Alignments were visualised using T-Coffee and BoxShade.[277]

8.10.5 Structure visualisation

Protein structures and 3D small molecules were visualised and figures prepared using UCSF Chimera.[32] Reaction schemes and 2D figures of small molecules were prepared using ChemDraw Professional 16.0 (PerkinElmer).

Chapter 9: Appendix

9.1 DNA sequences

Hexahistidine tags are underlined and TEV cleavage sites given in *italics*.

***Nn*NCS5**

CAT ATG CAT CAT CAC CAC CAT CAT TCT AGC GGC GTC GAC TTG GGT ACT GAG AAT CTG TAT TTT CAG AGC ATG ATT CAT AGC GTG AGC ACC GAA CTG GAA GTG GAT CTG CCG GCG GAT GAT ATT TGG GCG GTG TAT AGC AGC CCG GAA CTG CCG AAA CTG GTG GTG AAA CTG ATG CCG CAT GTG TAT GAT AAA ATT GAT ATT GTG GAA GGC GAT GGC GGC GTG GGC ACC GTG CTG CAG ATT GTG CTG ACC CCG GAA ATG ATG GAA CCG CGC ACC TGG AAA GAA AAA TTT GTG GAA ATT AAC GAT GGC CGC CGC AAA AAA GTG GTG CGC CAG ATT GAA GGC GGC TAT CTG GAT ATG GGC TTT CAT TTT TAT GAA GAT ATT TTT AAA ATT AAA AAA AAA AGC GAT AGC AGC TGC ATT ATT AAA AGC AAA AGC GTG TTT CGC GTG GAT CAT AAA CAT AAA GCG AAC GCG AGC CTG GTG ACC CCG GAT GCG AGC GCG GAA ATG GCG AAA GCG GTG GCG GAA TAT GCG AAA CAG AAA AAA GCG AAC AGC AGC AGC AGC AGC AAA GAT AAA GCG AAA GCG TGC TAT GAA TAA TCT CGA G

***Nn*NCS7**

CAT ATG CAT CAT CAC CAC CAT CAT TCT AGC GGC GTC GAC TTG GGT ACT GAG AAT CTG TAT TTT CAG AGC ATG ATG ACC GCG CGC GTG ACC AAC GAA ATG GAA GTG GGC GTG CCG GCG GAT GAT GTG TGG GCG GTG TAT GGC AGC CCG GAT CTG CCG AAA CTG TTT GTG CAG CTG ATG CCG CAG GTG TAT AAA CGC AAC GAT GTG CTG GAA GGC GAT GGC ACC GTG GGC ACC GTG ATT CTG ATT GAA CTG GAT GAT GCG CTG CCG GAA CCG CGC ATT TGG AAA GAA AAA TTT ATT AAA ATT GAT CAT CAG GAA CGC GAA AAA CTG GTG CGC GTG ATT GAA GGC GGC TTT CTG GAT ATT GGC TTT CGC AGC TTT GAT ATT ATT TTT AAA GTG ATT GAA AAA GAT GCG AGC AGC TGC ATT ATT CAG AGC ACC ACC GCG TTT GAA CTG GAT GAT AAA TTT GAA GAT AAC GCG AAC CGC ATT ACC GCG GGC ACC CTG TGG TGG GTG GCG AAA GCG ATT AGC AAC TAT GTG ATT CAG AAC AAA AGC AAA AGC AAA AGC GAT AAC AAC TAA TCT CGA G

SalSyn v1 (pTac promoter)

GGC AGT GAG CGC AAC GCA ATT AAA AAA GAT CAA AAA ATT TGA CAA TTA ATC ATC GGC TCG TAT AAT GTG TGA ATT GTG AGC GGA TAA CAA TTT GAT AAA TTG TTA AAG AGC AGT AGG AGA AAC ATA TGC AGA TTT TTG TGA AAA CCC TGA CCG GCA AAA CCA TTA CCC TGG AAG TGG AAC CGA GCG ATA CCA TTA AAA ACG TGA AAG CGA AAA TTC AGG ATA AAG AAG GCA TTC CGC CGG ATC AGC AGC GCC TGA TTT TTG CGG GCA AAC AGC TGG AAG ATG GCC GCA CCC TGA GCG ATT ATA ACA TTC AGA AAA AAA GCA CCC TGC ATC TGG TGC TGC GCC TGC GCG TGG ATT ATC ACC ATC ACC ATC ACC ATT AAT TGT TCA TTC AGA ACG CTC GGT CTT GCA CAC CGG GCG TTT TTT CTT TGT GAG TCC A

SalSyn v2 (pPl promoter)

GGC AGT GAG CGC AAC GCA ATT AAA AAA GAT CAA AAA ATG TTG ACA TAA ATA CCA CTG GCG GTT ATA CTG AGG AAT TGT GAG CGG ATA ACA ATT TGA TAA ATT GTT AAA GAG CAG TAG GAG AAA CAT ATG CAG ATT TTT GTG AAA ACC CTG ACC GGC AAA ACC ATT ACC CTG GAA GTG GAA CCG AGC GAT ACC ATT AAA AAC GTG AAA GCG AAA ATT CAG GAT AAA GAA GGC ATT CCG CCG GAT CAG CAG CGC CTG ATT TTT GCG GGC AAA CAG CTG GAA GAT GGC CGC ACC CTG AGC GAT TAT AAC ATT CAG AAA AAA AGC ACC CTG CAT CTG GTG CTG CGC CTG CGC GTG GAT TAT CAC CAT CAC CAT CAC CAT TAA TTG TTC ATT CAG AAC GCT CGG TCT TGC ACA CCG GGC GTT TTT TCT TTG TGA GTC CA

SalSyn v3 (T7 promoter)

GGC AGT GAG CGC AAC GCA ATT AAA AAA GAT CAA AAA ATT TAA TAC GAC TCA CAA TCG CGG AGC TGA GGG TGT GAA TTG TGA GCG GAT AAC AAT TTG ATA AAT TGT TAA AGA GCA GTA GGA GAA ACA TAT GCA GAT TTT TGT GAA AAC CCT GAC CGG CAA AAC CAT TAC CCT GGA AGT GGA ACC GAG CGA TAC CAT TAA AAA CGT GAA AGC GAA AAT TCA GGA TAA AGA AGG CAT TCC GCC GGA TCA GCA GCG CCT GAT TTT TGC GGG CAA ACA GCT GGA AGA TGG CCG CAC CCT GAG CGA TTA TAA CAT TCA GAA AAA AAG CAC CCT GCA TCT GGT GCT GCG CCT GCG CGT GGA TTA TCA CCA TCA CCA TCA CCA TTA ATT CTA GCA TAA CCC CTT GGG GCC TCT AAA CGG GTC TTG AGG GGT TTT TTG

9.2 Calibration curves

Calibration curves were prepared using products purified by preparative HPLC or from commercial sources. Samples prepared for standard curves of phenyethylamines were prepared with 0.5 molar equivalents of sodium ascorbate (rt = 1.8 min) to minimise oxidation of the amine. Absorbance is given in mAU at 280 nm. Representative analytical HPLC traces of the isolated products are shown, with analysis performed using analytical HPLC method 1.

9.2.1 Chapter 2: Novel Biocatalytic Routes Towards Tetrahydroisoquinoline Alkaloids

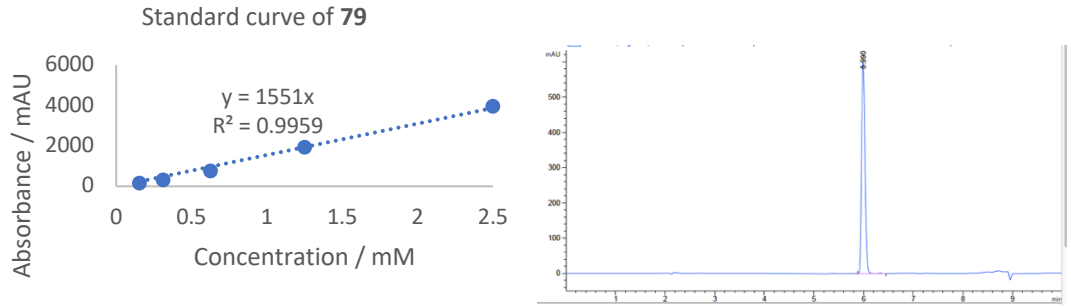


Figure 9.1: Calibration curve and HPLC spectrum of 79.

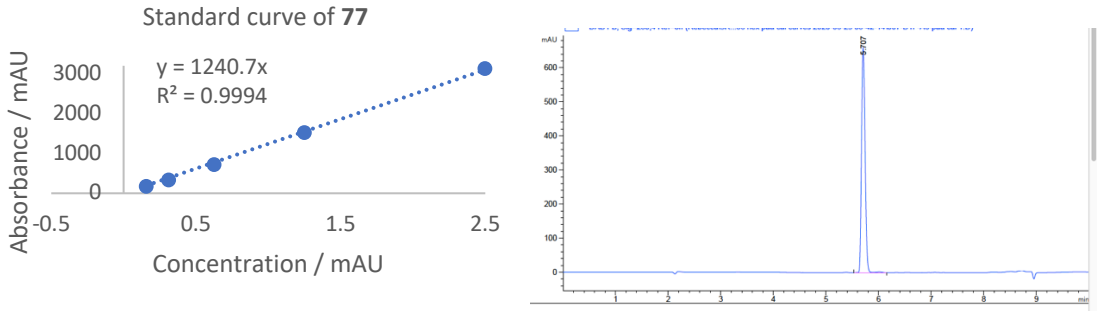


Figure 9.2: Calibration curve and HPLC spectrum of 77.

9.2.2 Chapter 3: The Acceptance and Kinetic Resolution of α–Methyl Substituted Aldehydes by Norcoclaurine Synthases

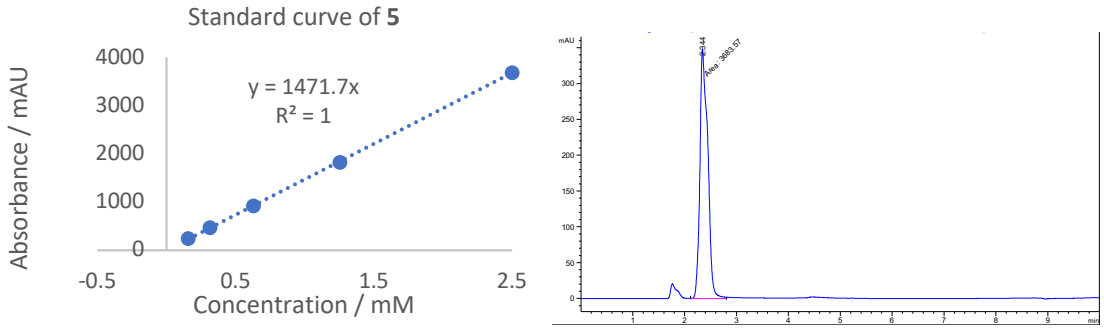


Figure 9.3: Calibration curve and HPLC spectrum of 5.

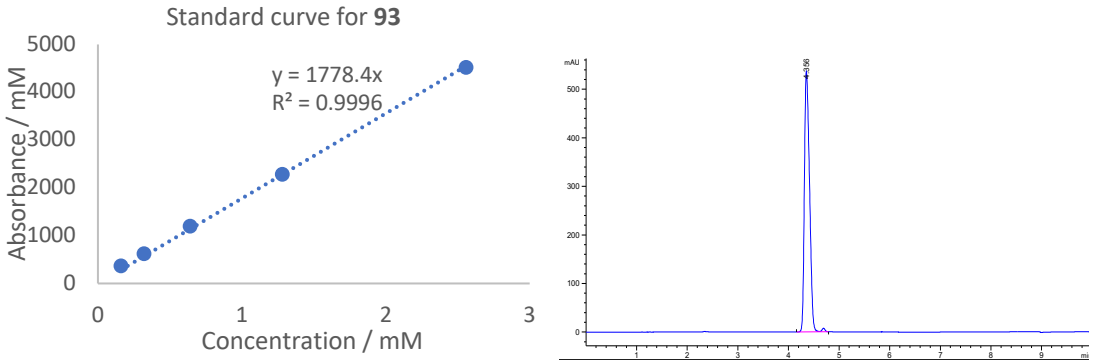


Figure 9.4: Calibration curve and HPLC spectrum of 93.

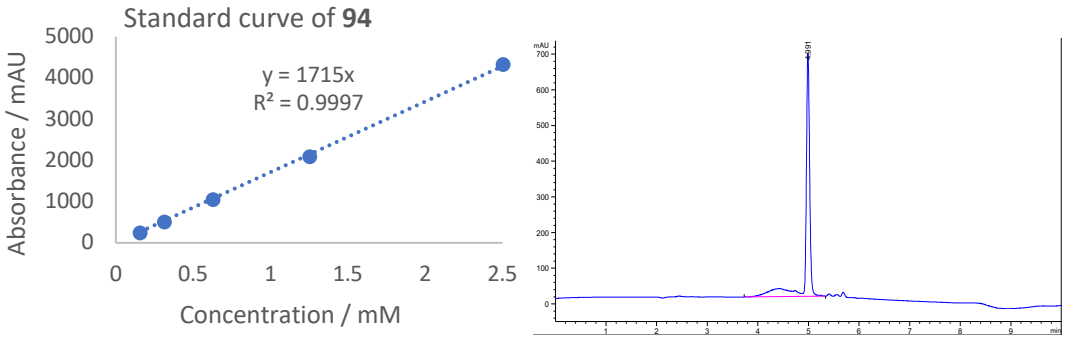


Figure 9.5: Calibration curve and HPLC spectrum of 94.

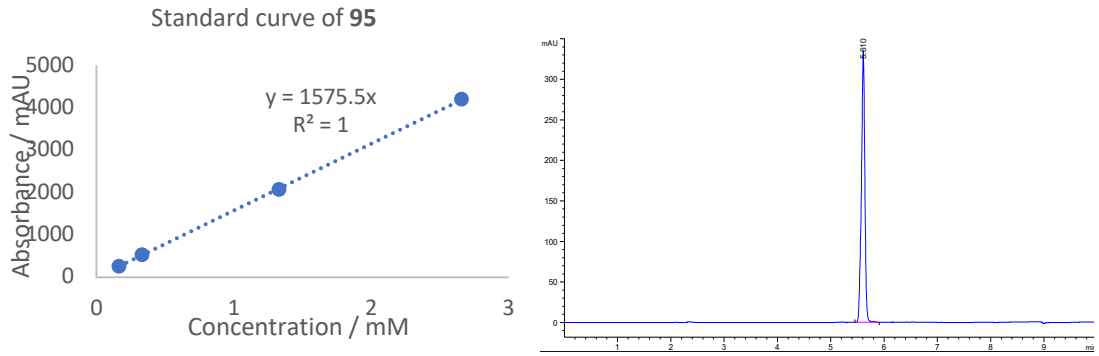


Figure 9.6: Calibration curve and HPLC spectrum of 95.

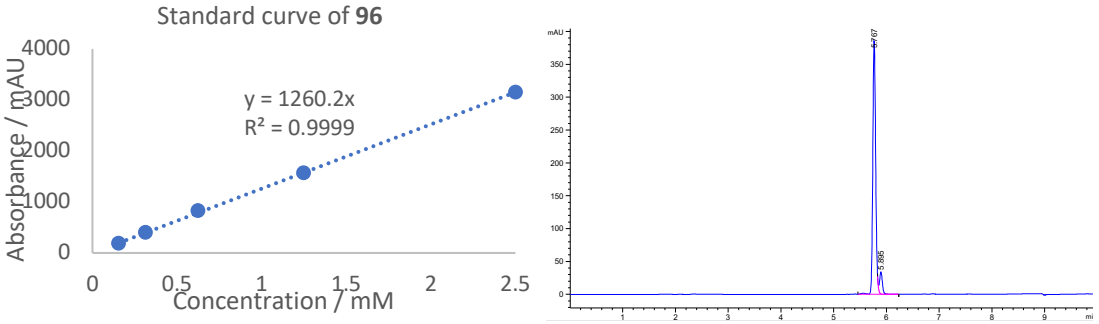


Figure 9.7: Calibration curve and HPLC spectrum of 96.

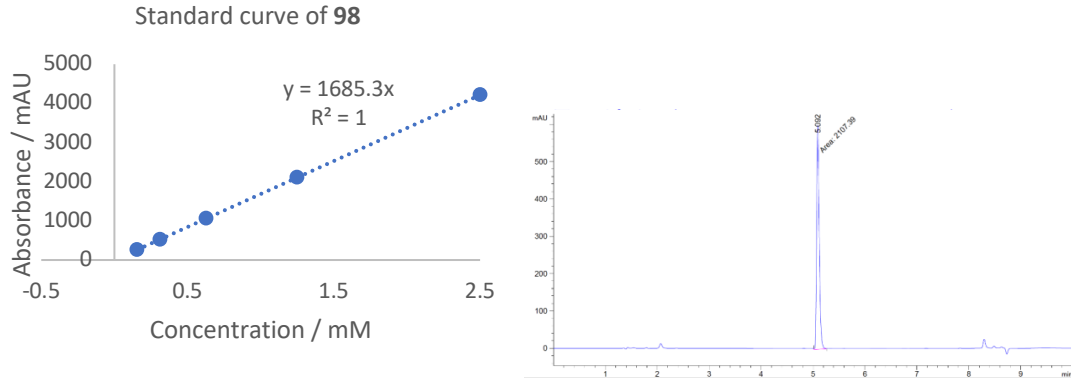


Figure 9.8: Calibration curve and HPLC spectrum of 98.

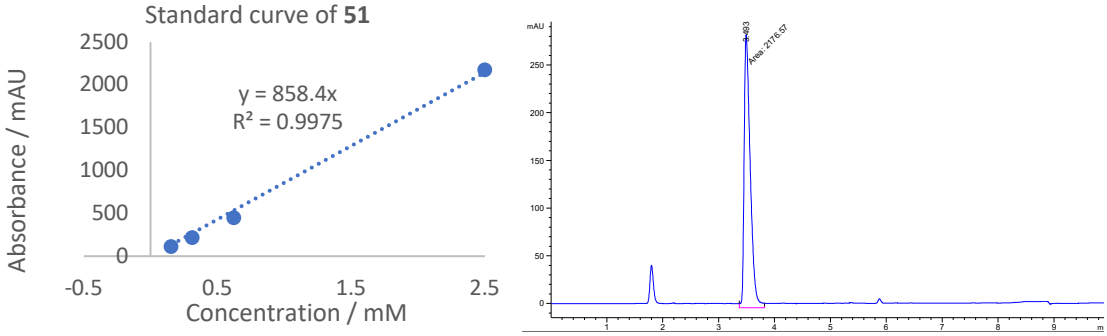


Figure 9.10: Calibration curve and HPLC spectrum of 51.

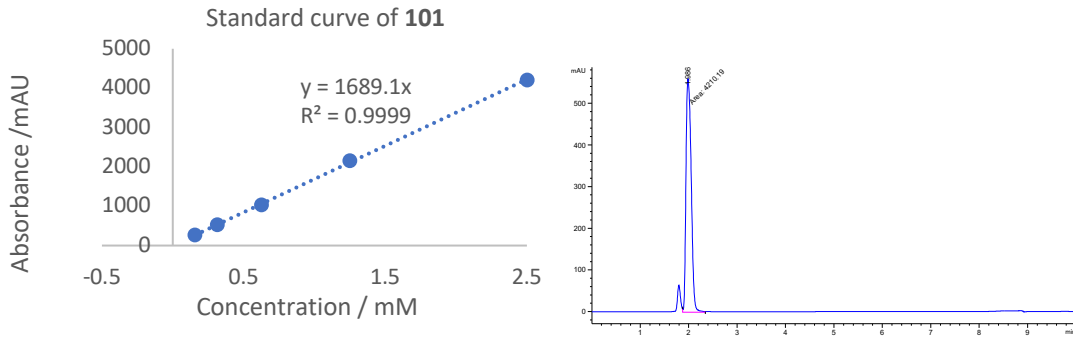


Figure 9.11: Calibration curve and HPLC spectrum of 101.

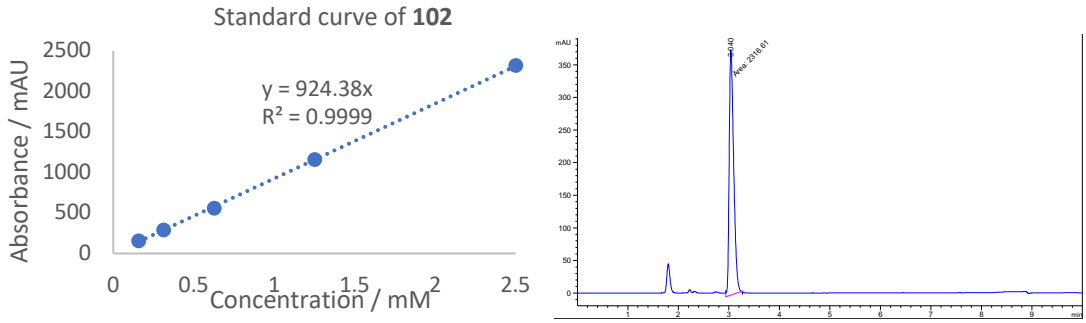


Figure 9.12: Calibration curve and HPLC spectrum of 102.

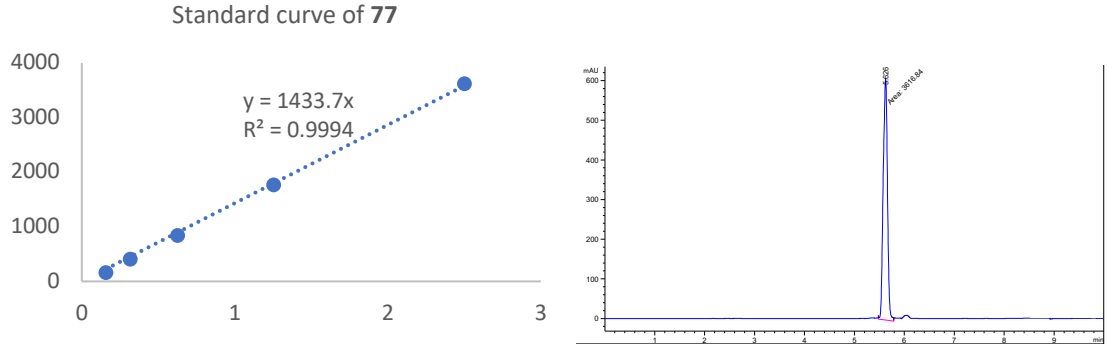


Figure 9.13: Calibration curve and HPLC spectrum of 77.

9.2.3 Chapter 4: Chemoenzymatic Cascades Towards Methylated Tetrahydroprotoberberine and Protoberberine Alkaloids

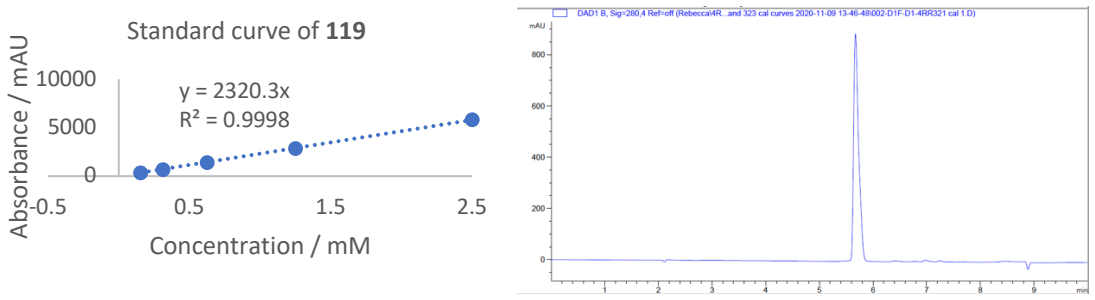


Figure 9.14: Calibration curve and HPLC spectrum of 119.

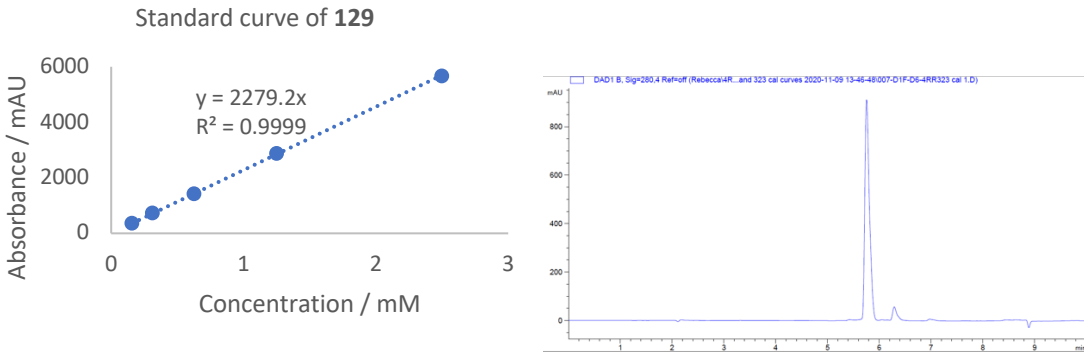


Figure 9.15: Calibration curve and HPLC spectrum of 129.

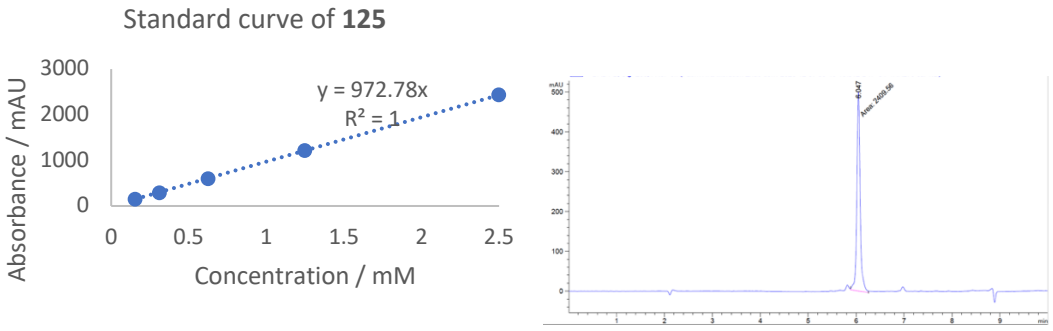


Figure 9.16: Calibration curve and HPLC spectrum of 125.

9.2.4 Chapter 5: Single Step Syntheses of (1*S*)-Aryl-Tetrahydroisoquinolines by Norcoclaurine Synthases

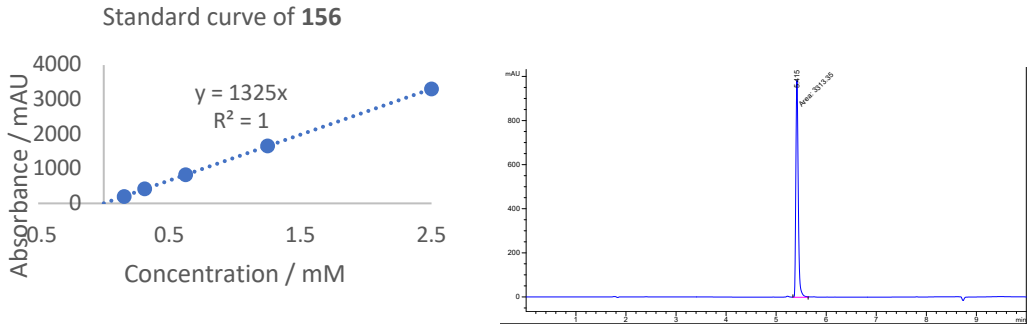


Figure 9.17: Calibration curve and HPLC spectrum of 156.

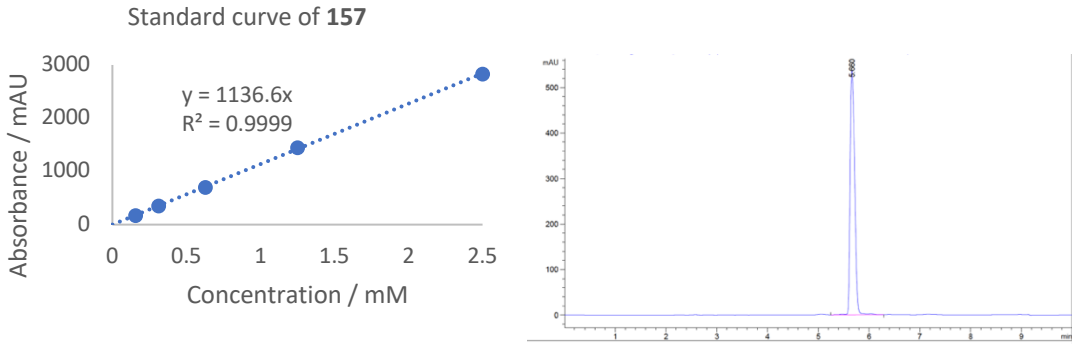


Figure 9.18: Calibration curve and HPLC spectrum of 157.

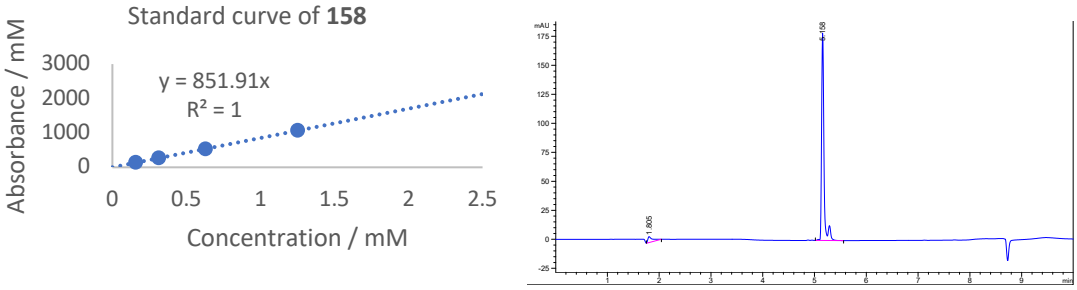


Figure 9.19: Calibration curve and HPLC spectrum of 158.

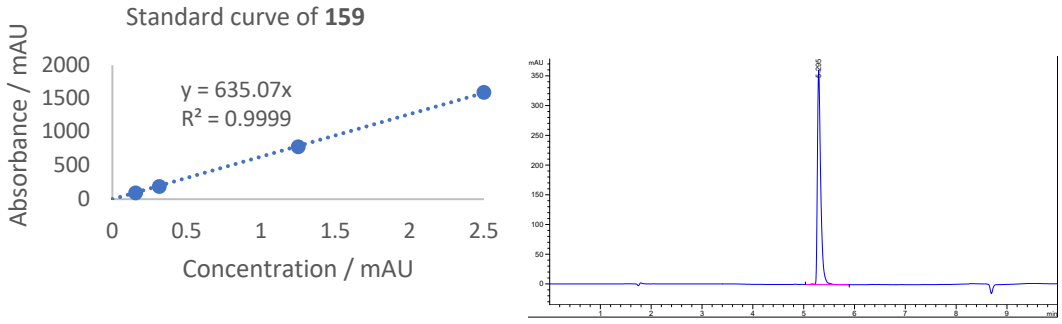


Figure 9.20: Calibration curve and HPLC spectrum of 159.

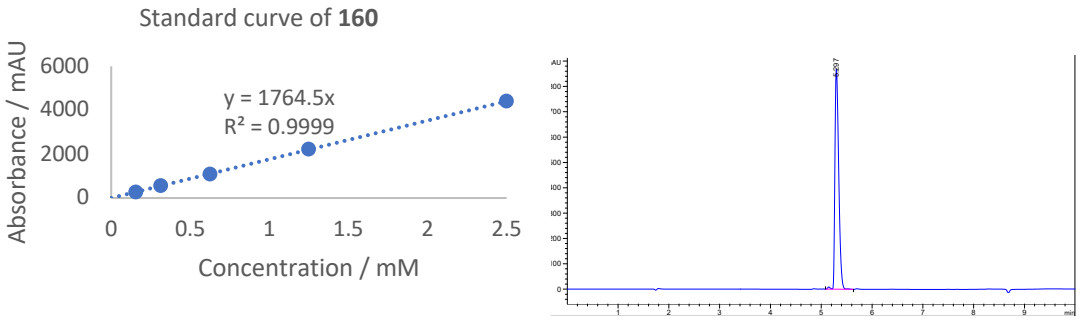


Figure 9.21: Calibration curve and HPLC spectrum of 160.

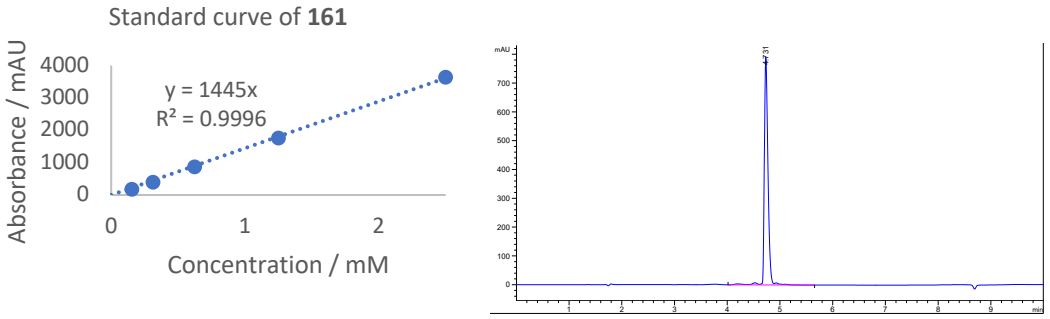


Figure 9.22: Calibration curve and HPLC spectrum of 161.

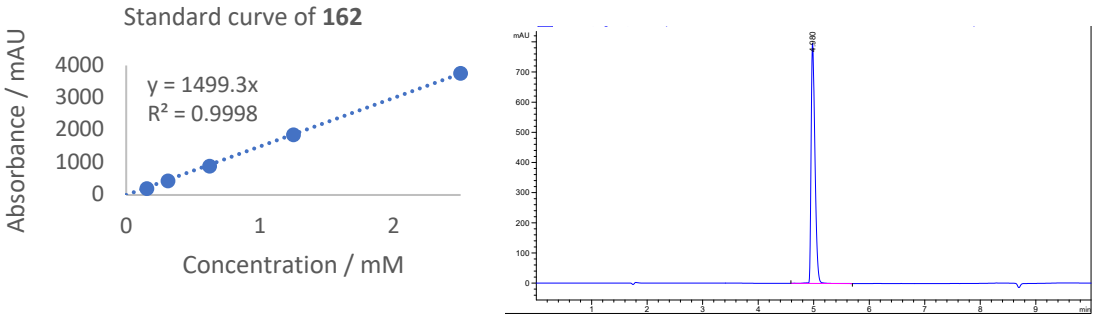


Figure 9.23: Calibration curve and HPLC spectrum of 162.

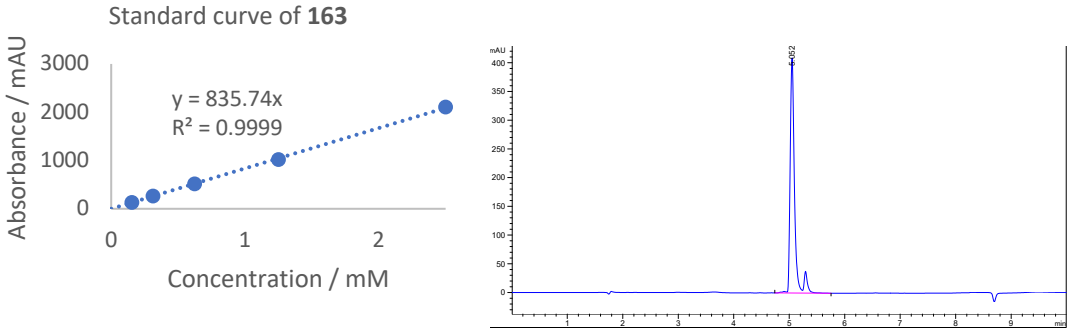


Figure 9.24: Calibration curve and HPLC spectrum of 163.

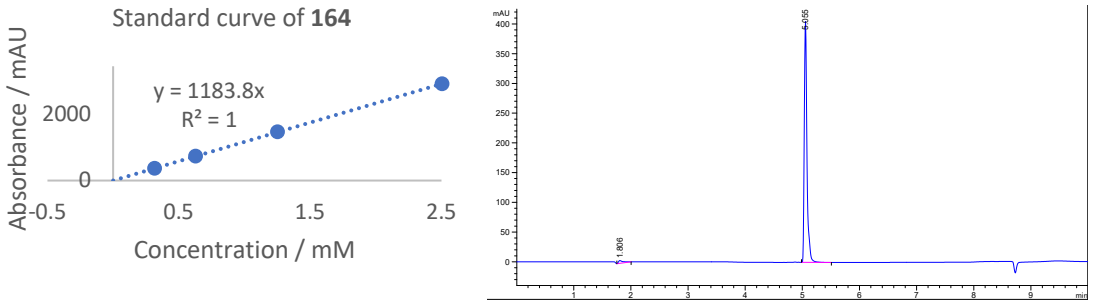


Figure 9.25: Calibration curve and HPLC spectrum of 164.

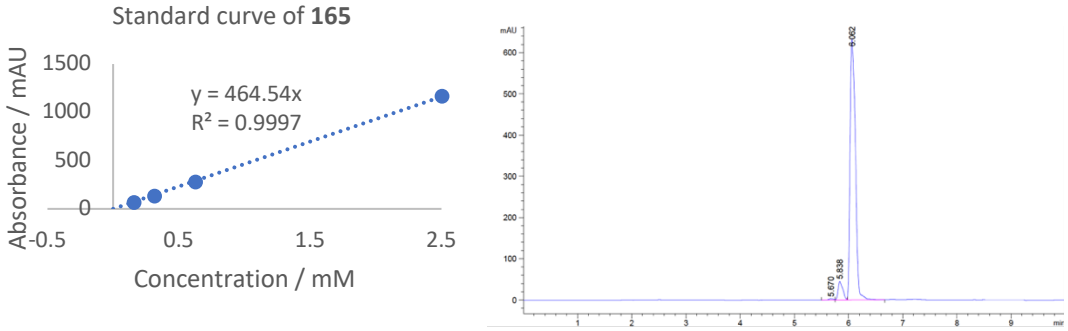


Figure 9.26: Calibration curve and HPLC spectrum of 165.

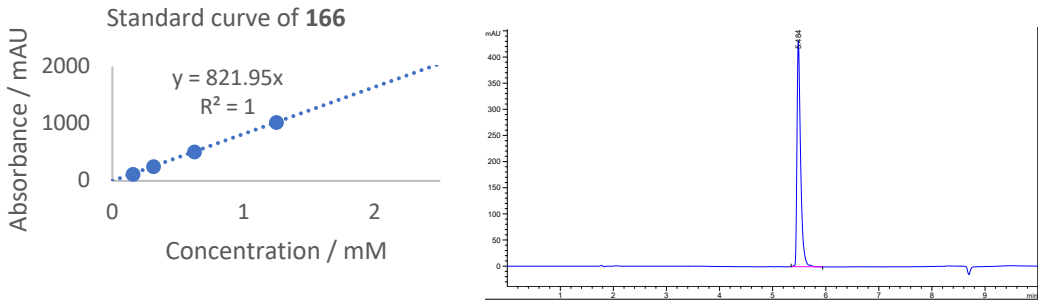


Figure 9.27: Calibration curve and HPLC spectrum of 166.

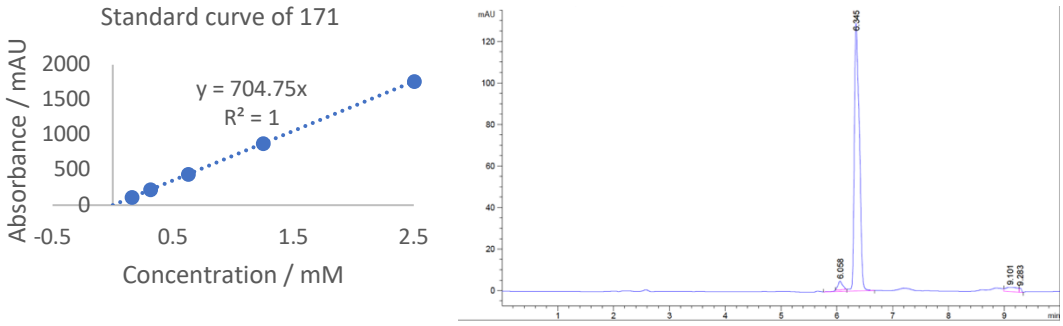


Figure 9.28: Calibration curve and HPLC spectrum of 171.

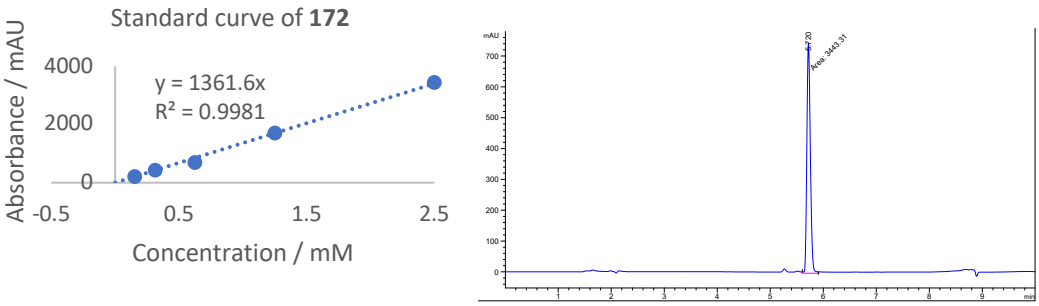


Figure 9.29: Calibration curve and HPLC spectrum of 172.

9.2.4 Chapter 6: *N*-methyl-phenethylamines as substrates for norcoclaurine synthases

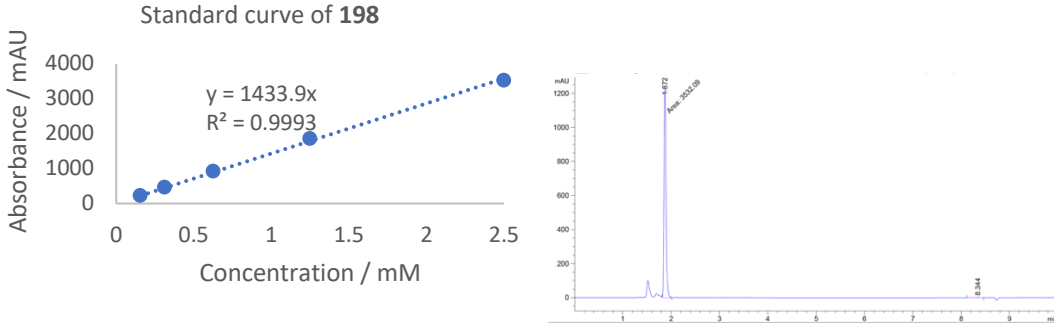


Figure 9.30: Calibration curve and HPLC spectrum of 198.

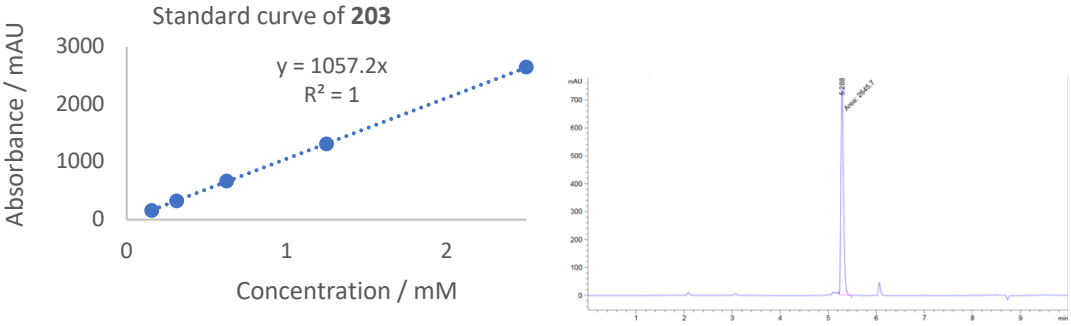


Figure 9.31: Calibration curve and HPLC spectrum of 203.

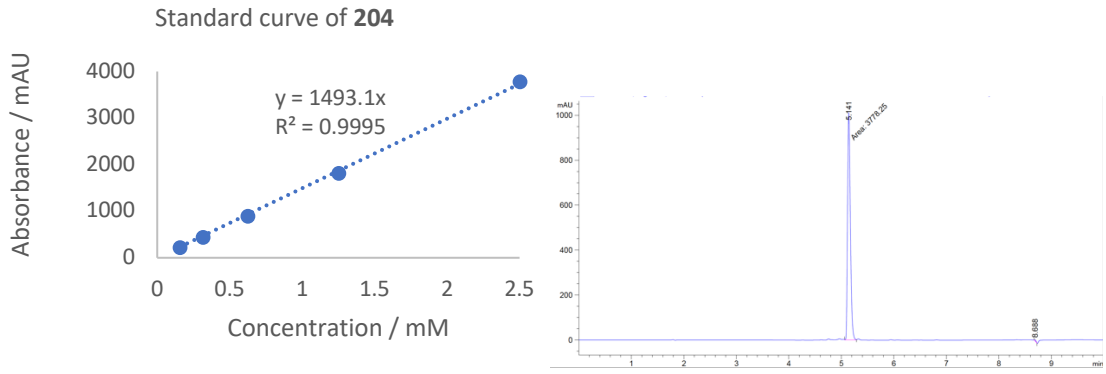


Figure 9.32: Calibration curve and HPLC spectrum of 204.

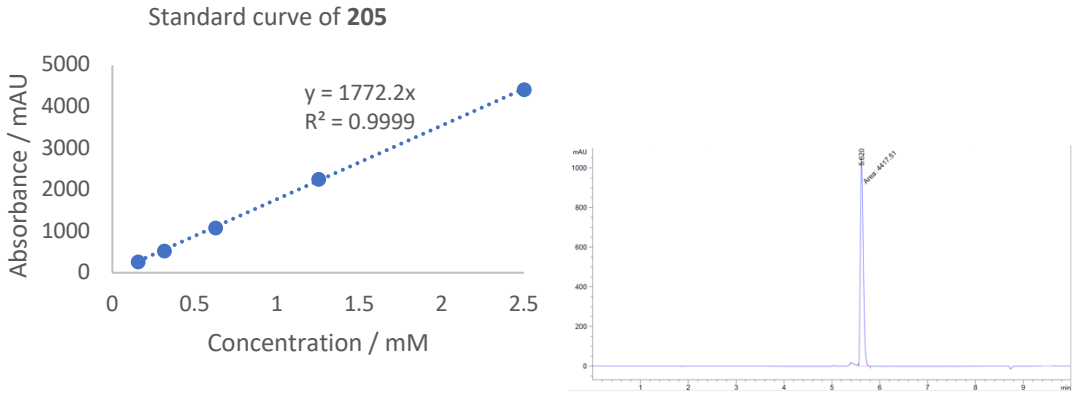


Figure 9.33: Calibration curve and HPLC spectrum of 205.

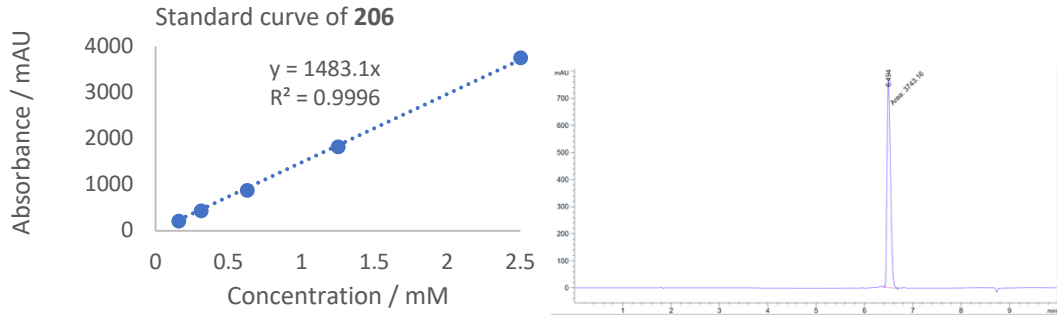


Figure 9.34: Calibration curve and HPLC spectrum of 206.

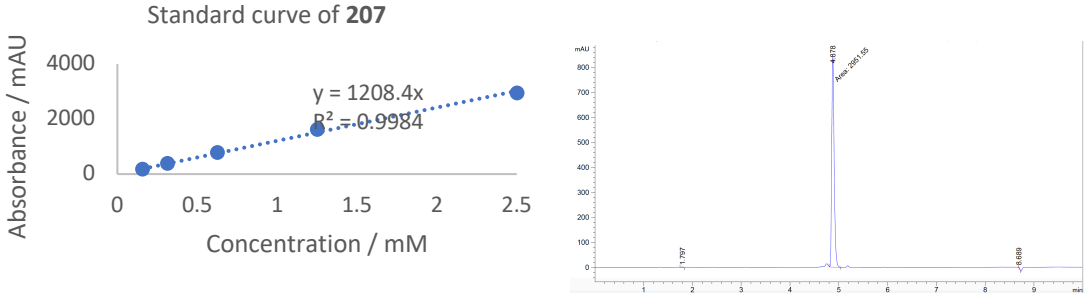


Figure 9.35: Calibration curve and HPLC spectrum of 207.

9.3 Chiral HPLC analyses

9.3.1 Chapter 2: Novel Biocatalytic Routes Towards Tetrahydroisoquinoline Alkaloids

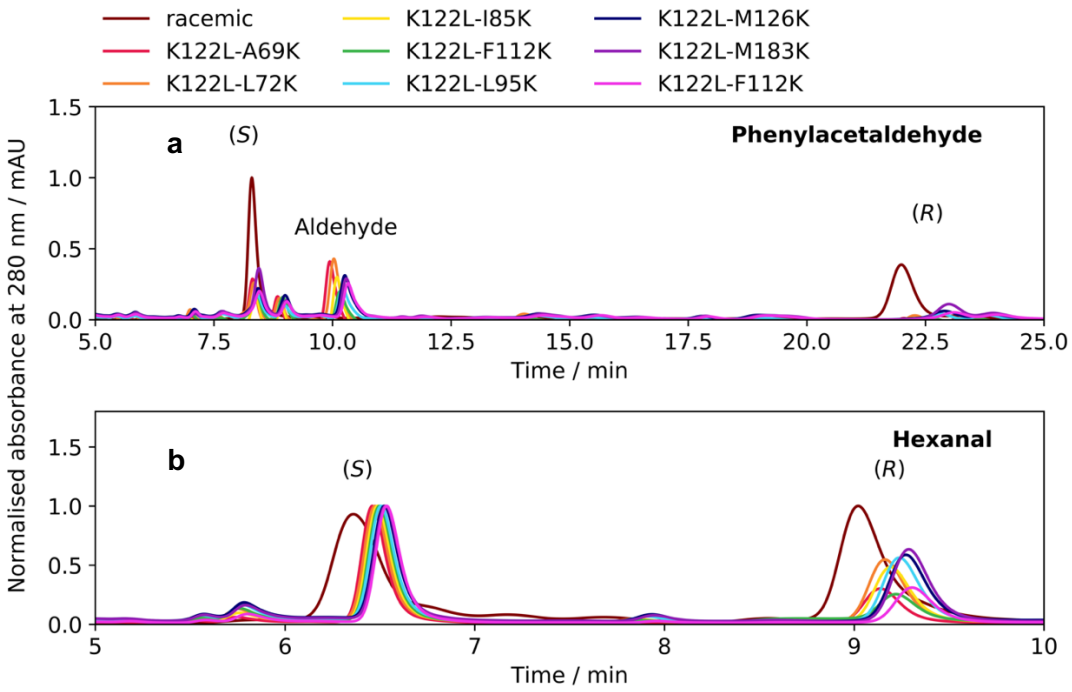


Figure 9.36: Chiral HPLC analysis of THIQ products, 77 and 79, generated by reactions with double point variants of Δ29*Tf*NCS. Reactions were performed in standard conditions using 1 mg mL^{-1} final concentration of enzyme using **a**. phenylacetaldehyde (**76**) or **b**. hexanal (**78**) as the aldehyde substrate. Samples were prepared using workup method 2 and analysis was performed using chiral HPLC method 5.

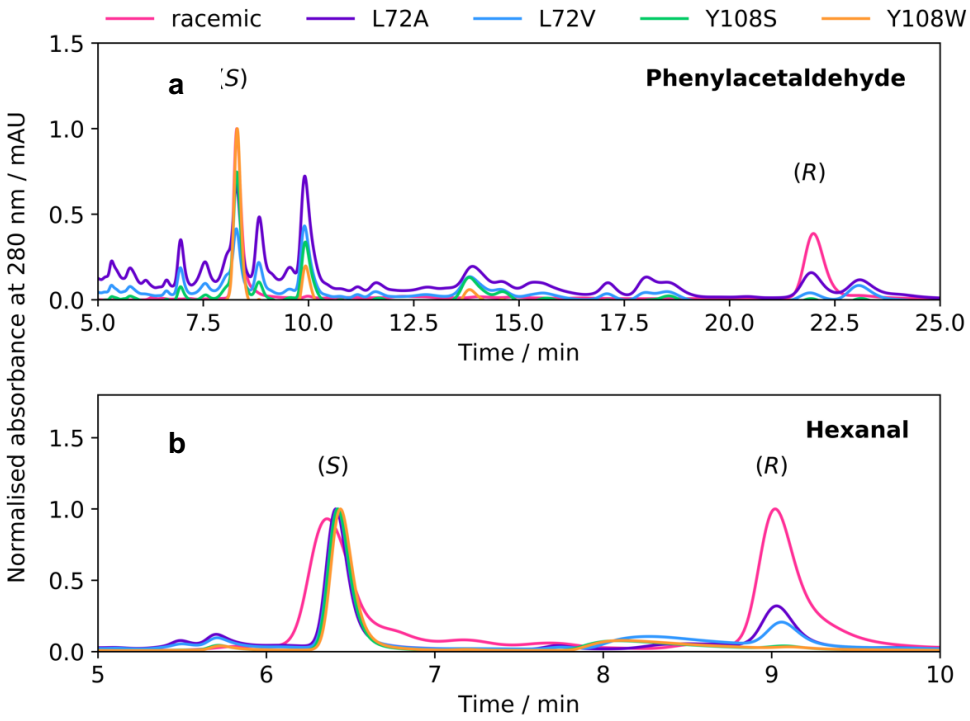


Figure 9.37: Chiral HPLC analysis of THIQ products, 77 and 79, generated by reactions with L72 and F80 variants of Δ33*Tf*NCS. Reactions were performed in standard conditions using 1 mg mL^{-1} final concentration of enzyme using **a**. phenylacetaldehyde (**76**) or **b**. hexanal (**78**) as the aldehyde substrate. Samples were prepared using workup method 2 and analysis was performed using chiral HPLC method 5.

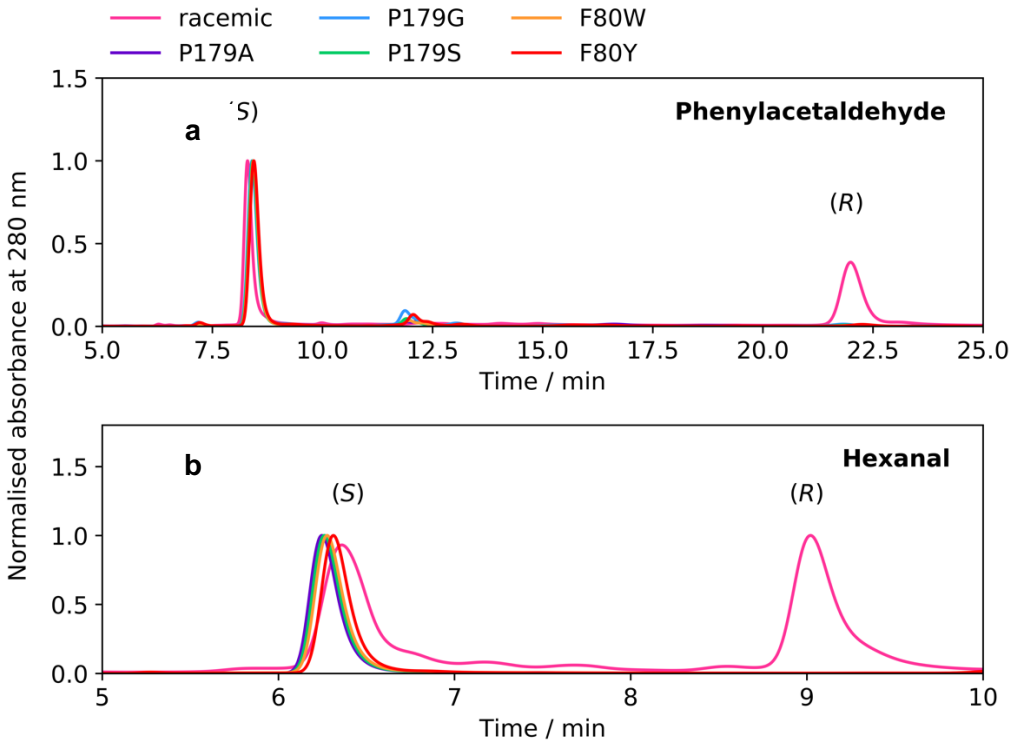


Figure 9.38: Chiral HPLC analysis of THIQ products, 77 and 79, generated by reactions with P179 and F80 variants of Δ33*Tf*NCS. Reactions were performed in standard conditions using 0.1 mg mL^{-1} final concentration of enzyme using **a**. phenylacetaldehyde (**76**) or **b**. hexanal (**78**) as the aldehyde substrate. Samples were prepared using workup method 2 and analysis was performed using chiral HPLC method 5.

9.3.2. Chapter 5: Single Step Syntheses of (1*S*)-Aryl-Tetrahydroisoquinolines by Norcoclaurine Synthases

For the chiral HPLC analyses of products **156**-**166**, traces shown are for reactions performed using 0.2 mg mL^{-1} final concentration of enzyme with a reaction time of 3 h. Acetonitrile was used as the reaction co-solvent. Samples were prepared for chiral HPLC analysis using workup method 2 and chiral HPLC method 5 was used unless otherwise stated. KPi = racemic standard, generated by a phosphate mediated Pictet-Spengler reaction using conditions described in the methods section and the product isolated by the same preparative HPLC method as the NCS-generated products. Various NCSs were used as the reaction catalyst: wild-type = WT, Δ 29*Tf*NCS-M97F = M97F, Δ29*Tf*NCS-M97V = M97V, Δ29*Tf*NCS-L76V = L76V. *Preparation of dopamine*

sample: A solution of dopamine HCl (5 mM) and sodium ascorbate (5 mM) was prepared.

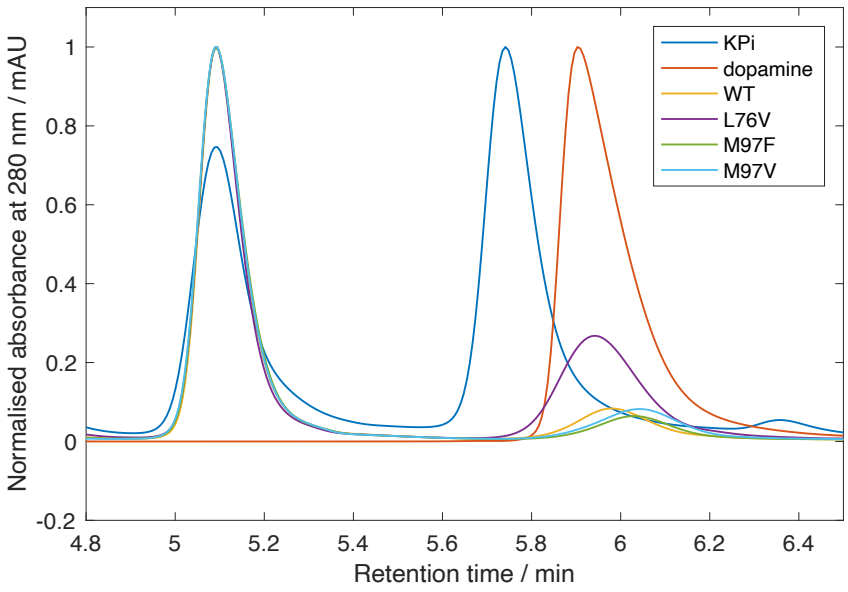


Figure 9.39: Chiral HPLC analysis of 156. Retention time (t_R) (major, *S* enantiomer) = 5.1 min, t_R (minor, *R* enantiomer) = 5.7 min. Samples were prepared using workup method 1 and analysis performed using chiral HPLC method 4.

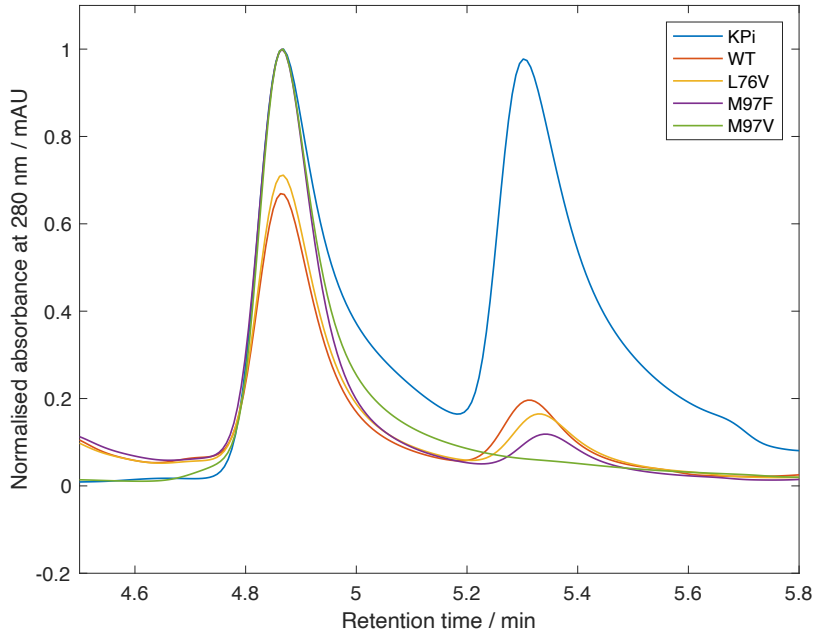


Figure 9.40: Chiral HPLC analysis of 157. t_R (major, *S* enantiomer) = 4.9 min, t_R (minor, *R* enantiomer) = 5.3 min. Samples were prepared using workup method 1 and analysis performed using chiral HPLC method 4.

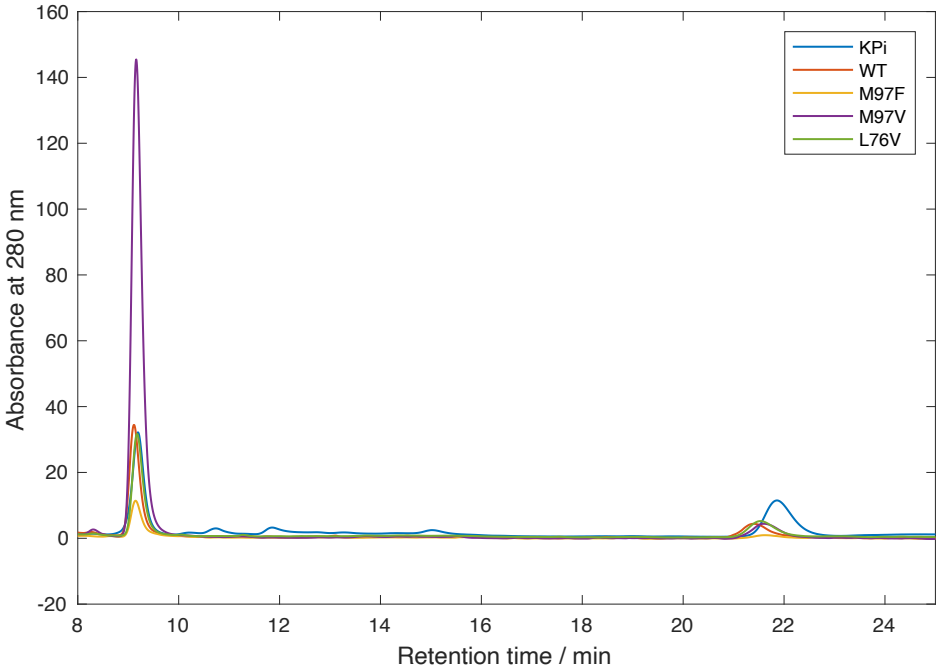


Figure 9.41: Chiral HPLC analysis of 158. t_R (major, *S* enantiomer) = 9 min, t_R (minor, *R* enantiomer) = 22 min.

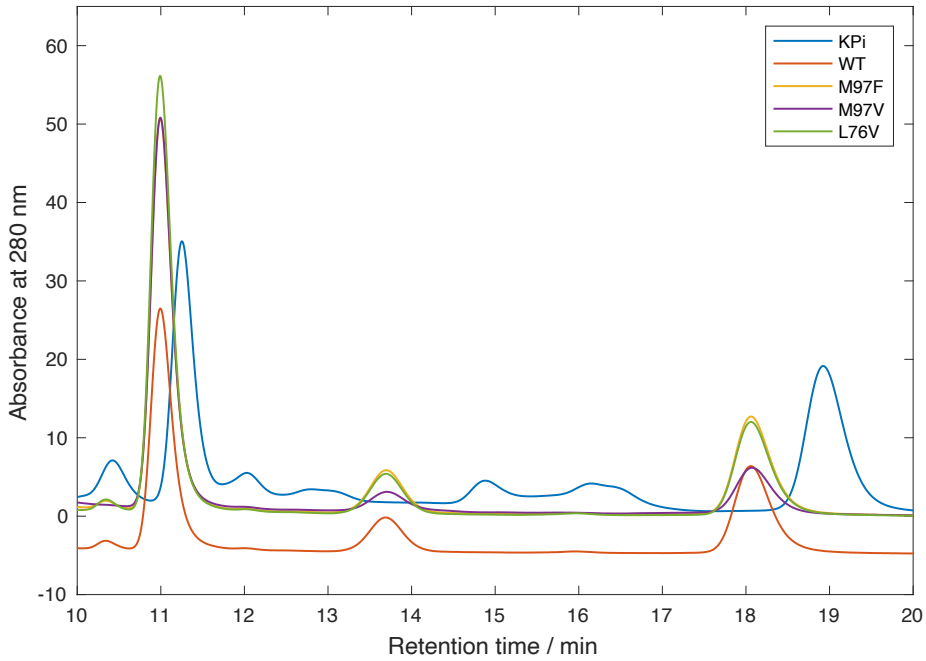


Figure 9.42: Chiral HPLC analysis of 159. t_R (major, *S* enantiomer) = 11 min, t_R (minor, *R* enantiomer) = 18-19 min.

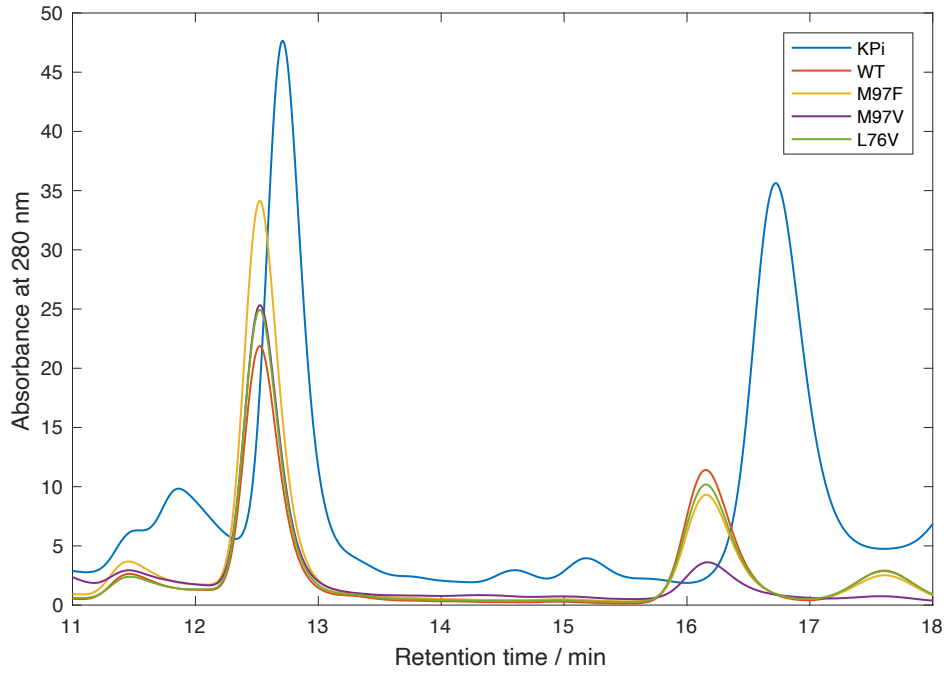


Figure 9.43: Chiral HPLC analysis of 160. t_R (major, *S* enantiomer) = 12-13 min, t_R (minor, *R* enantiomer) = 16 – 17 min.

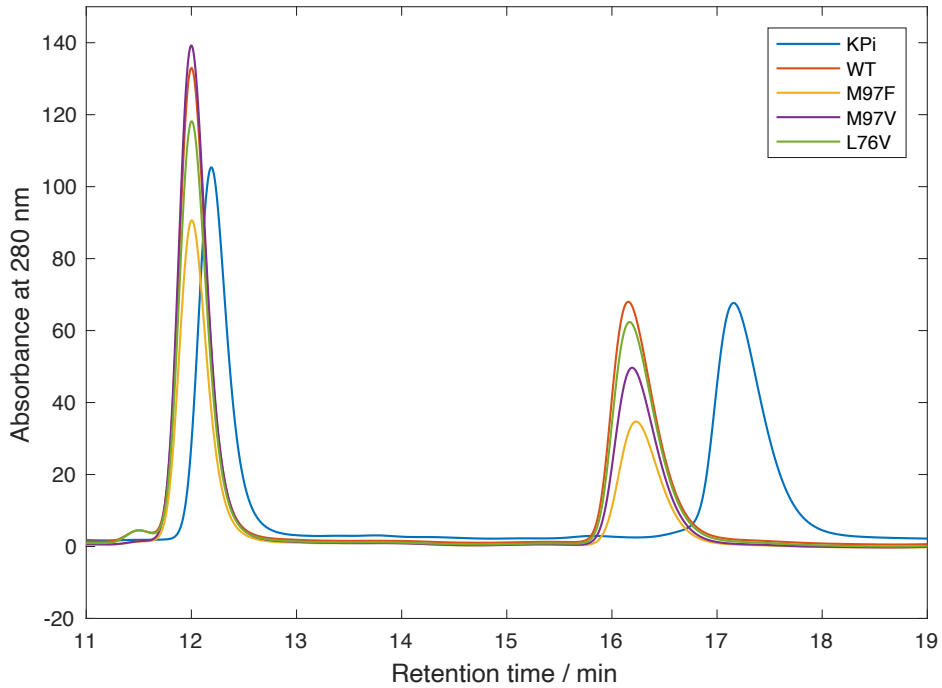


Figure 9.44: Chiral HPLC analysis of 161. t_R (major, *S* enantiomer) = 11 min, t_R (minor, *R* enantiomer) = 16-17 min.

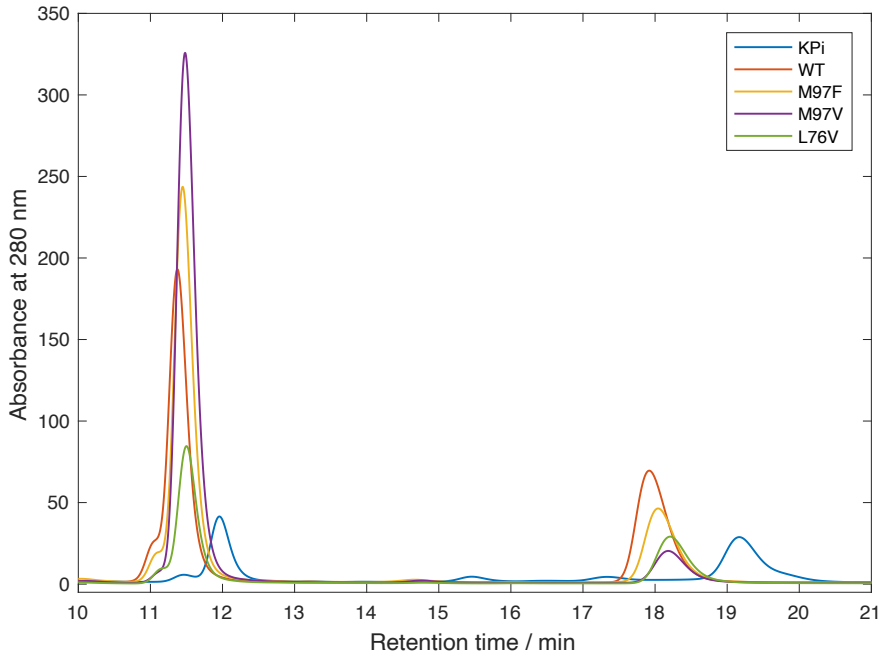


Figure 9.45: Chiral HPLC analysis of 162. t_R (major, *S* enantiomer) = 11 min, t_R (minor, *R* enantiomer) = 17-20 min.

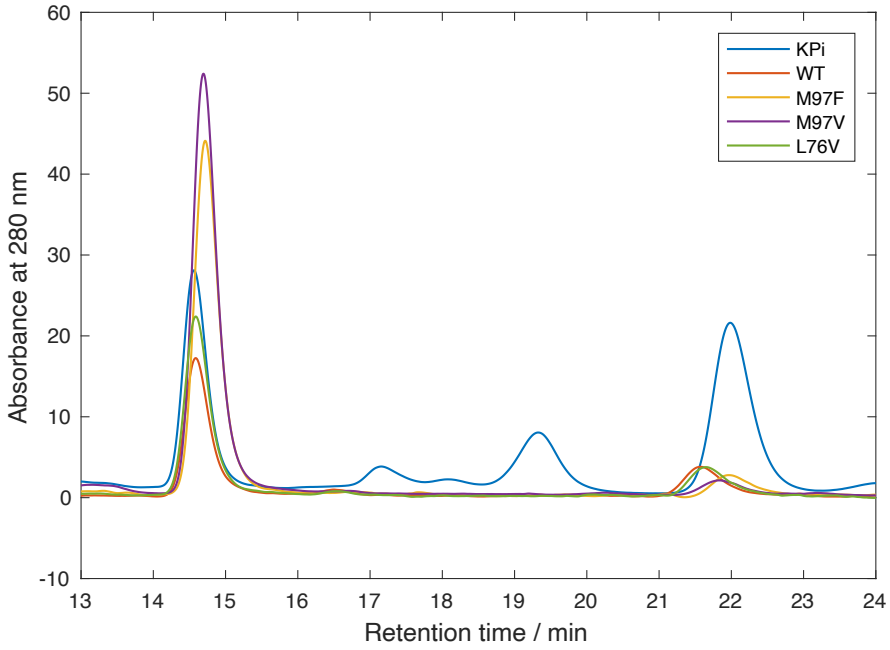


Figure 9.46: Chiral HPLC analysis of 163. t_R (major, *S* enantiomer) = 14-15 min, t_R (minor, *R* enantiomer) = 21-23 min.

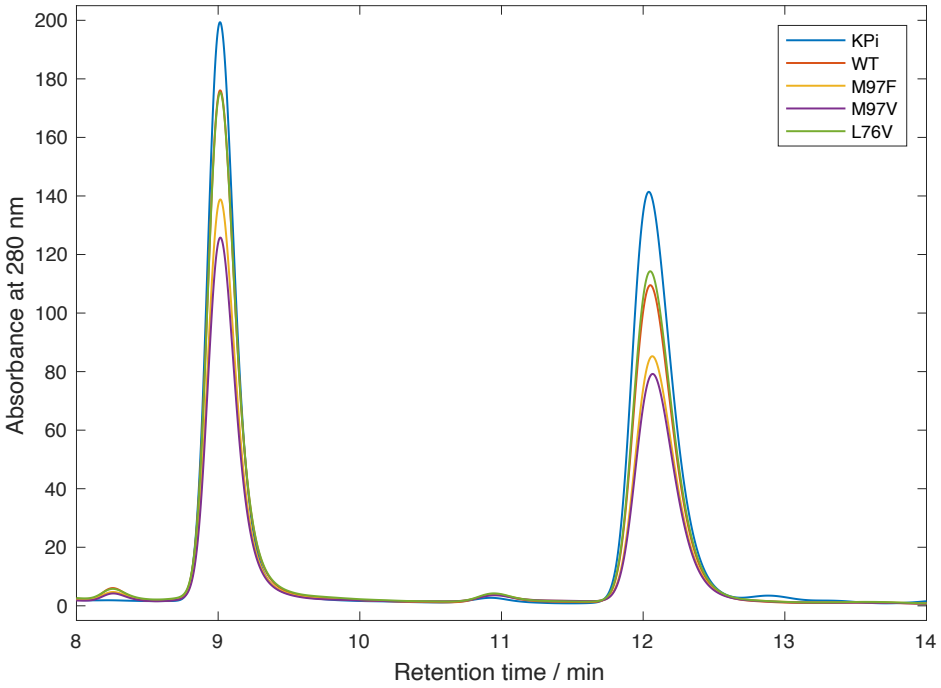


Figure 9.47: Chiral HPLC analysis of 164. t_R (major, *S* enantiomer) = 9 min, t_R (minor, *R* enantiomer) = 12 min.

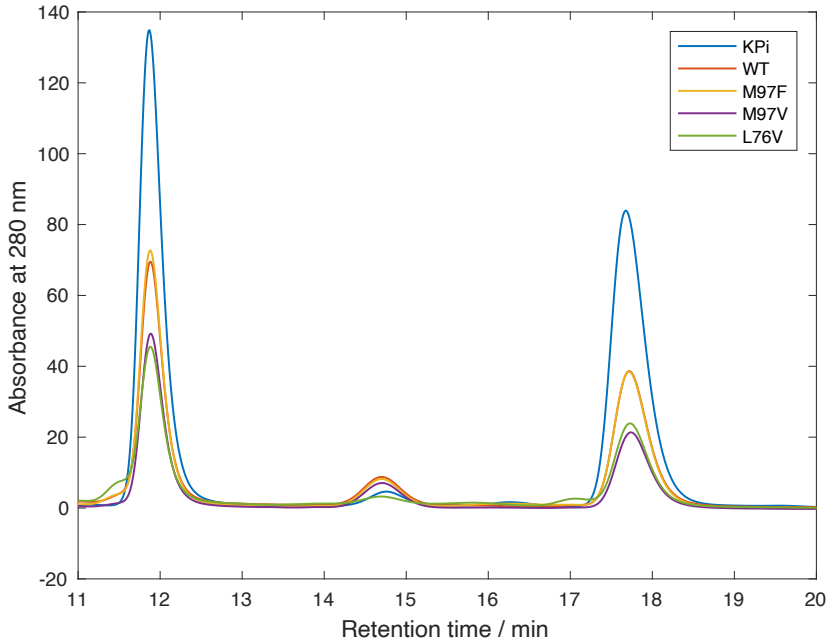


Figure 9.48: Chiral HPLC analysis of 165. t_R (major, *S* enantiomer) = 12 min, t_R (minor, *R* enantiomer) = 17-18 min.

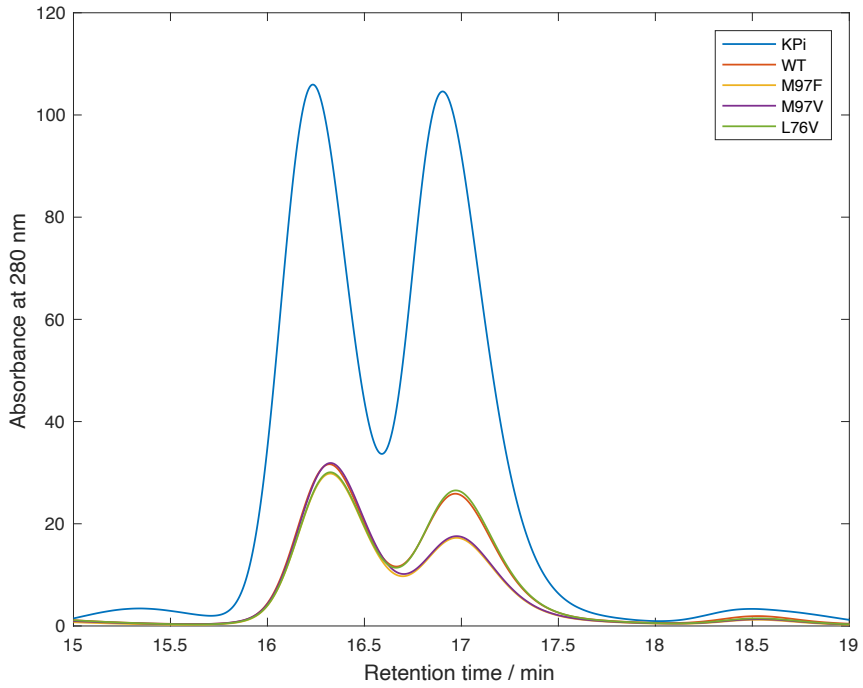


Figure 9.49: Chiral HPLC analysis of 166. t_R (major, *S* enantiomer) = 16 min, t_R (minor, *R* enantiomer) = 17 min.

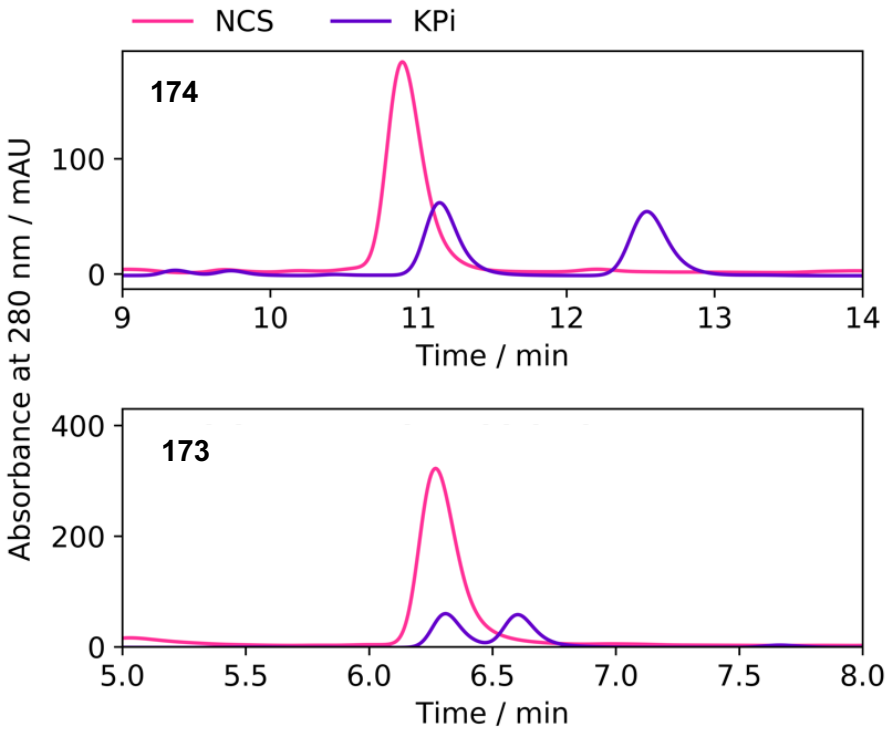


Figure 9.50: Chiral HPLC analyses of 173 and 174.

9.3.3 Chapter 6: *N*-methyl-Phenyethylamines as Substrates for Norcoclaurine Synthases

Chiral HPLC analyses shown are for reactions using standard conditions. 0.2 mg mL^{-1} final concentration of WT-Δ29*Tf*NCS was used with a reaction time of 18 h. The THIQ products (**203**, **204**, **206** and **207**) were isolated by preparative HPLC, any remaining TFA salts were removed by workup method 3 and chiral HPLC analysis was performed using method 5. Samples were prepared at 2.5 mM in the mobile phase and a 10 μL injection volume was used. Racemic standards (denoted racemic standard or KPi) were generated using a KPI-mediated reaction.

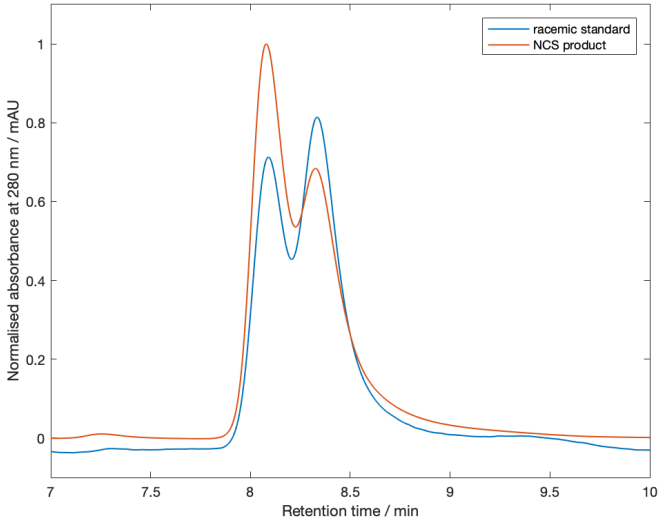


Figure 9.51: **Chiral HPLC analysis of 203.**

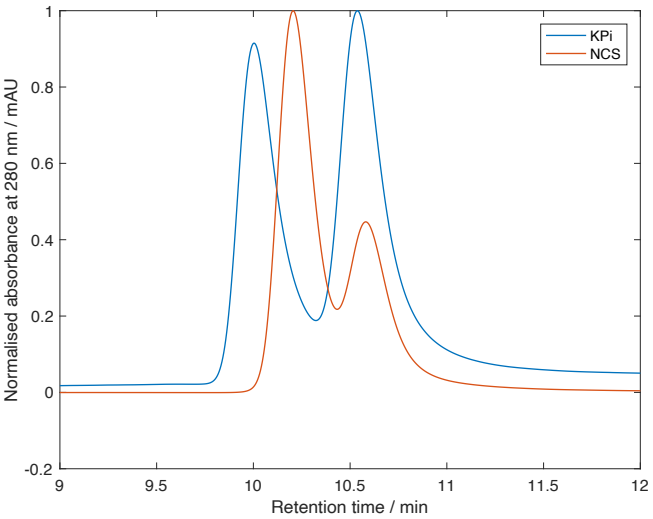


Figure 9.52: **Chiral HPLC analysis of 204.**

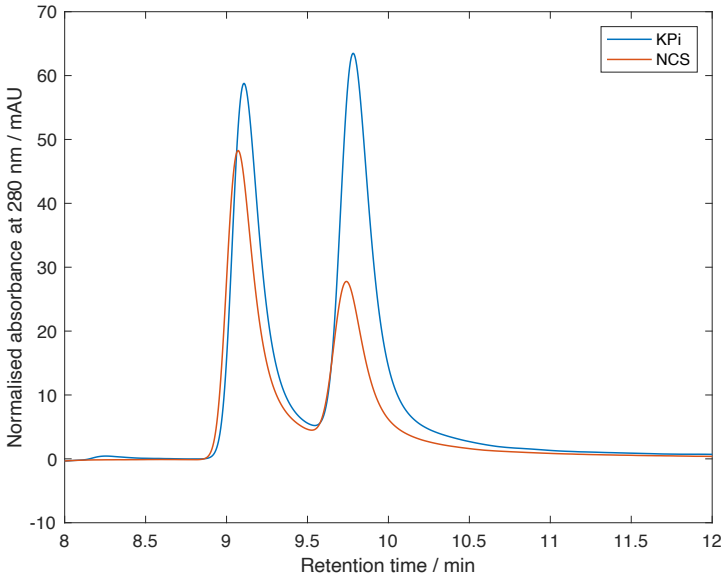


Figure 9.53: **Chiral HPLC analysis of 206 THIQ product.**

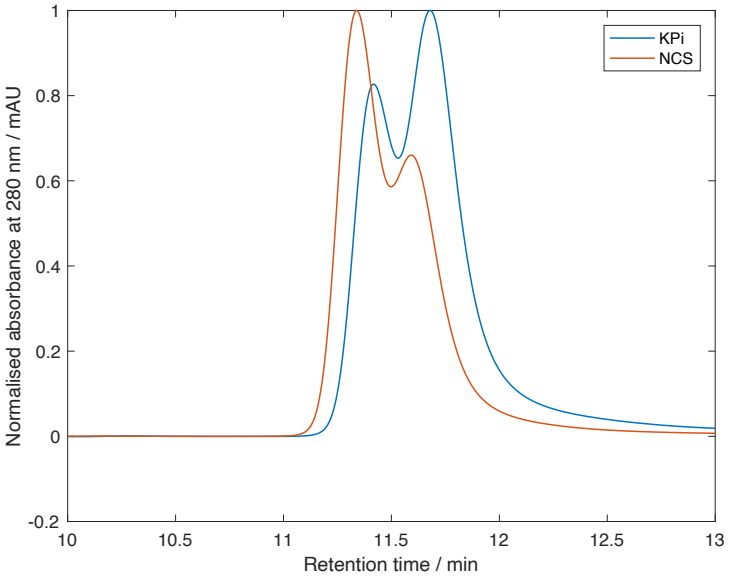


Figure 9.54: **Chiral HPLC analysis of 207 THIQ product.**

9.4 GCMS analysis of (*R*)-2-methylbutanol

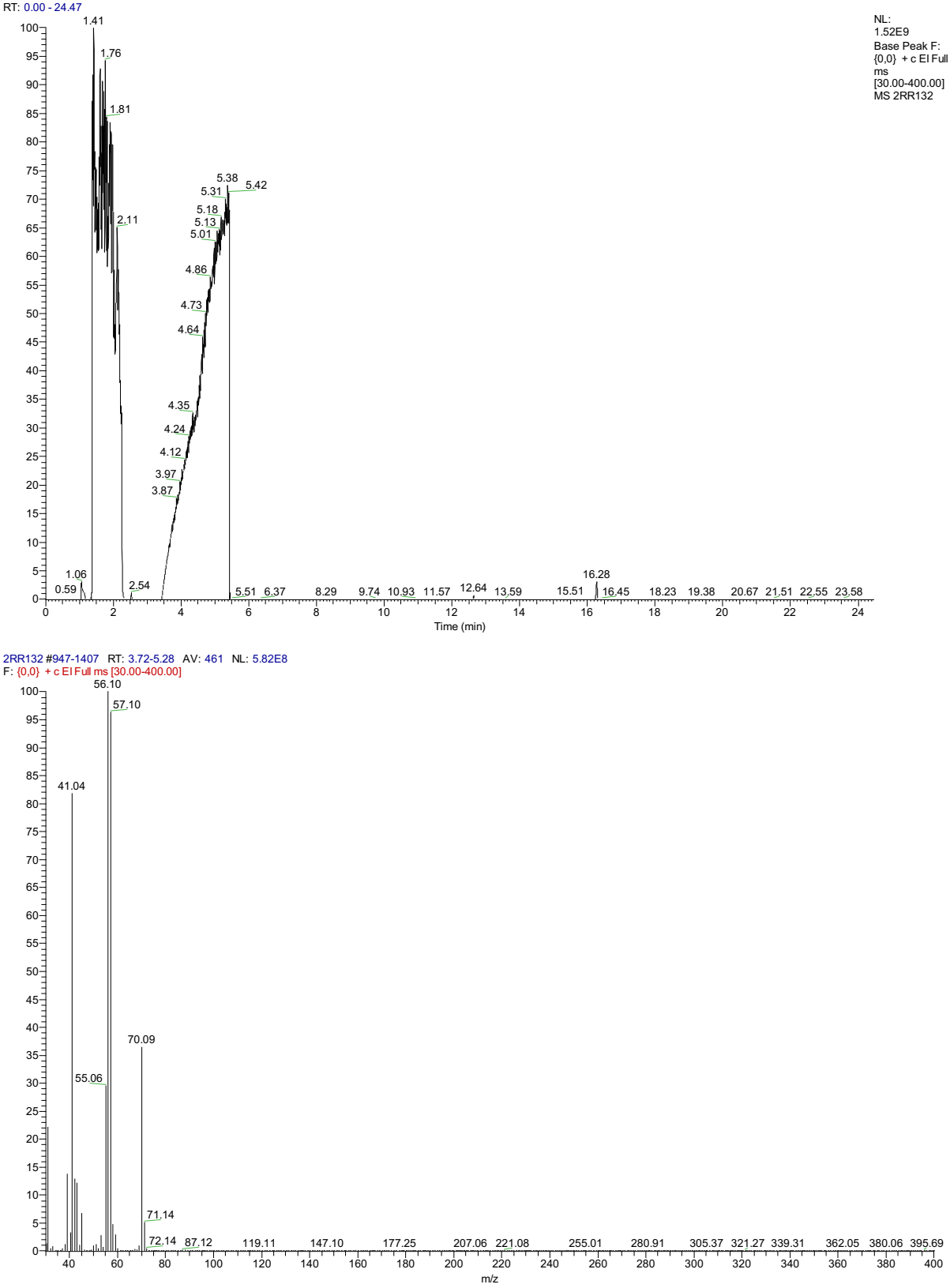
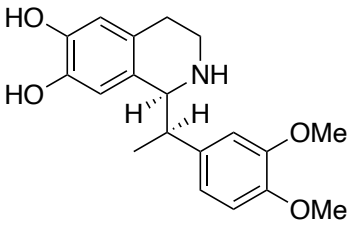
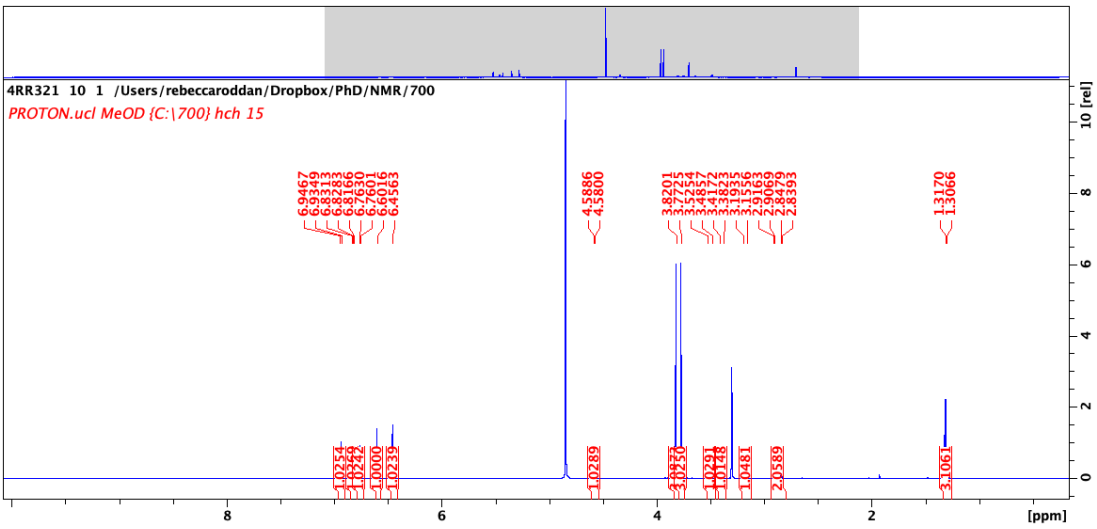


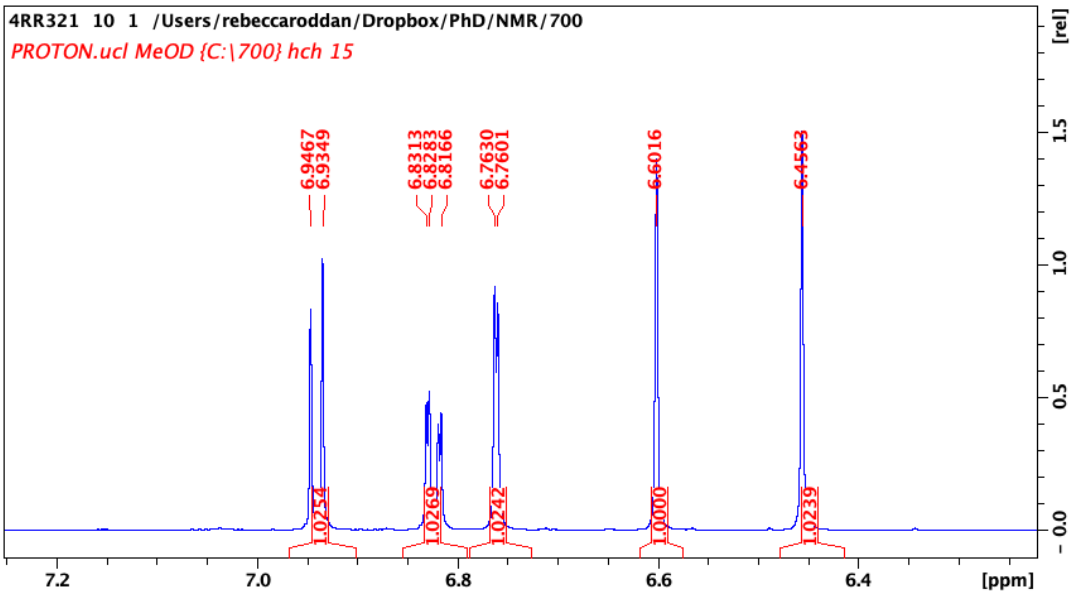
Figure 9.55: GCMS analysis of 103.

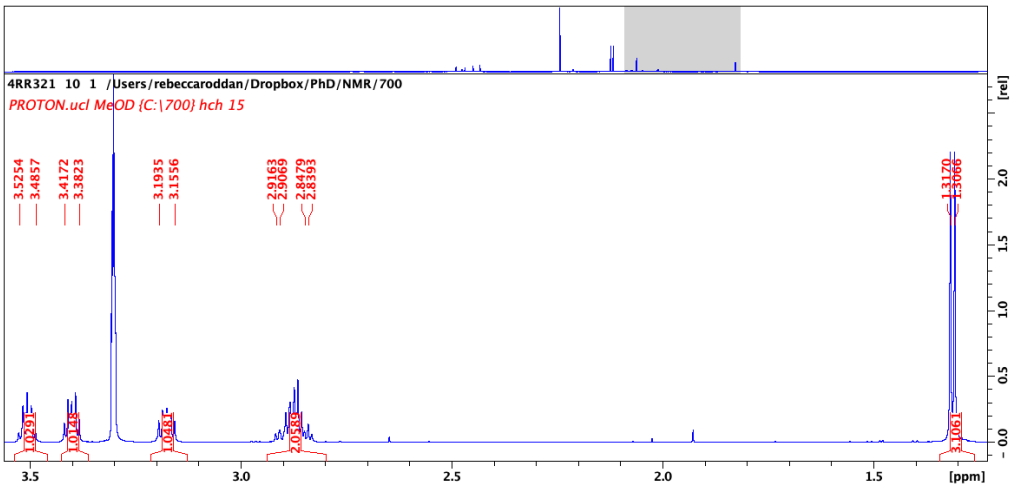
9.5 NMR spectra of compounds generated by work in Chapter 4

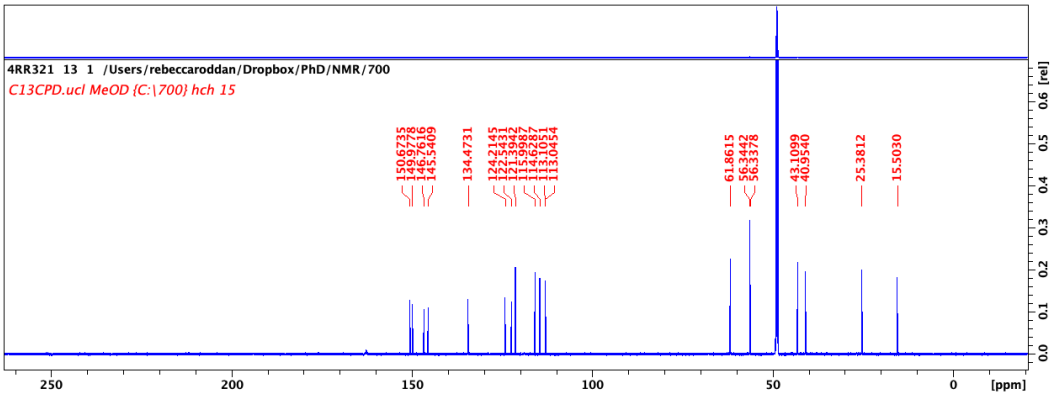
(*S*)-1-((*R*)-1-(3,4-Dimethoxyphenyl)ethyl)-1,2,3,4-tetrahydroisoquinoline-6,7-diol.TFA (199)

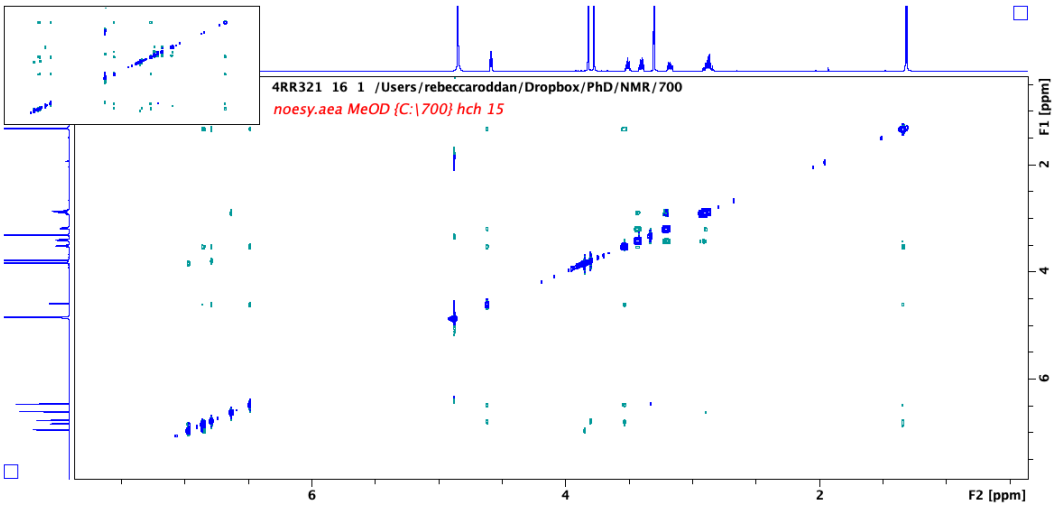




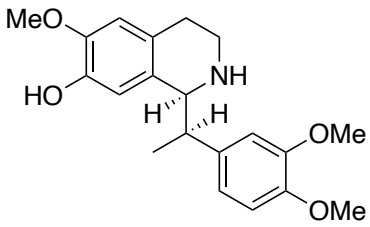


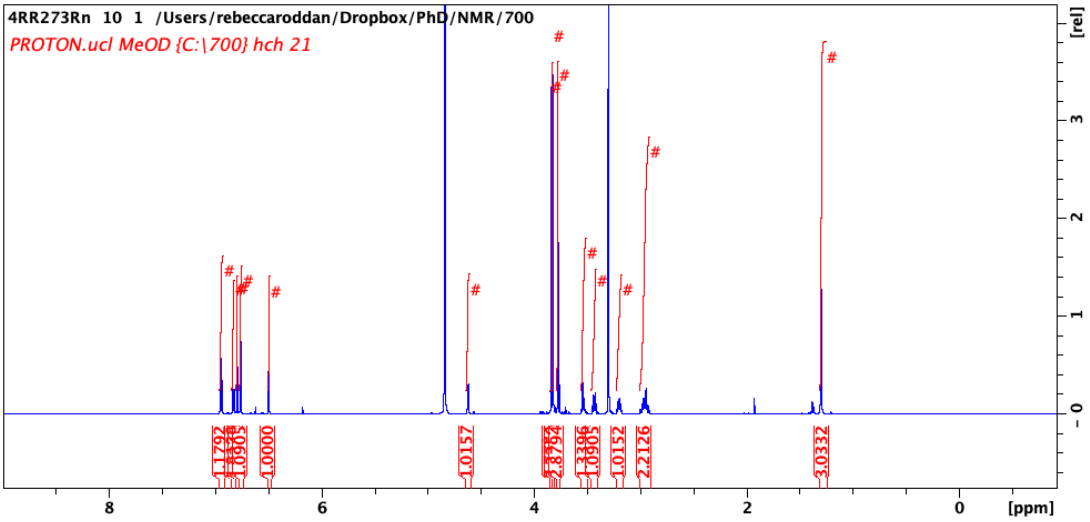

4RR321 10 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
PROTON.ucl MeOD {C:\700} hch 15
3.5254
3.4857
3.4172
3.3823
3.1935
3.1556
2.9163
2.9069
2.8479
2.8393
1.3170
1.3068
1.0291
1.0148
1.0281
2.0589
3.1061
[rel]
[ppm]

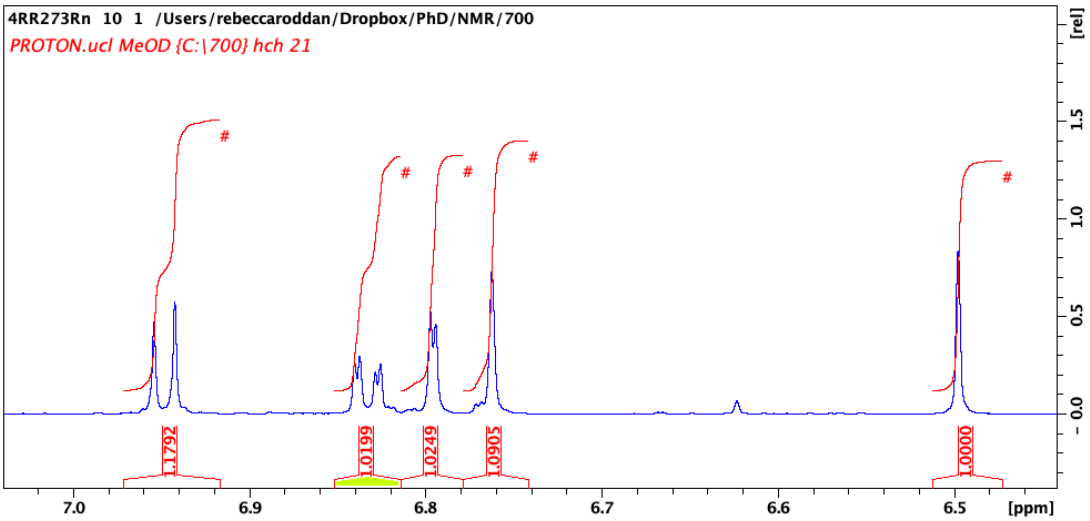

4RR321 13 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
C13CPD.ucl MeOD {C:\700} hch 15
150.6735
149.8778
146.7616
145.5409
134.4731
124.2145
122.5431
121.3942
114.6287
113.1051
113.0454
61.8615
56.3442
56.3378
43.1090
40.9540
25.3812
15.5030
[rel]
[ppm]

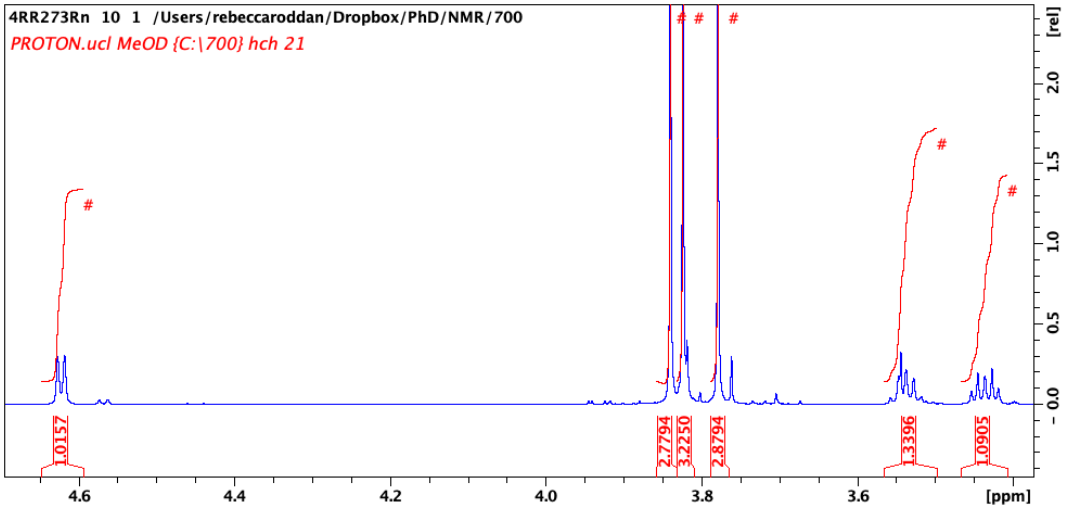

4RR321 16 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
noesy.aea MeOD {C:\700} hch 15
F1 [ppm]
F2 [ppm]

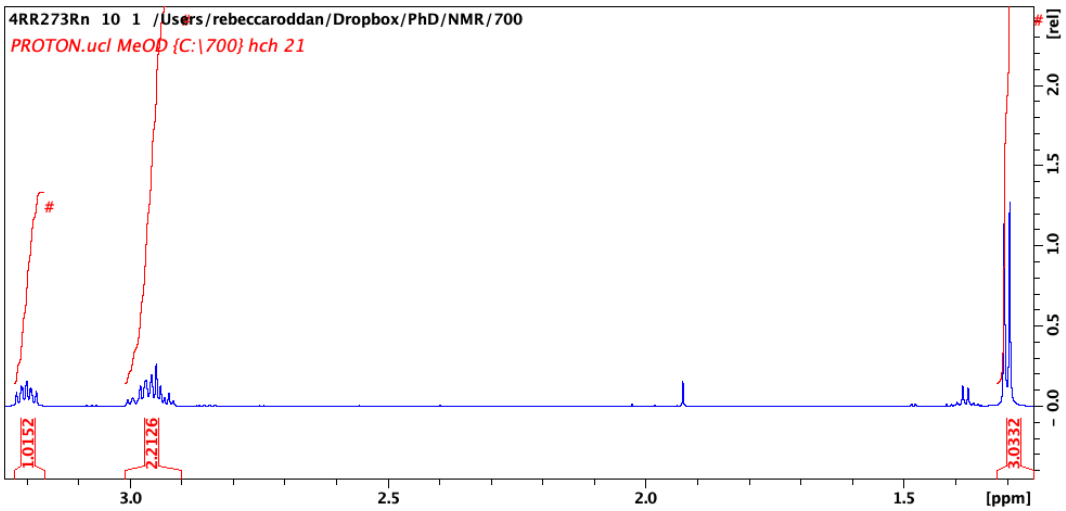
(*S*)-1-((*R*)-1-(3,4-Dimethoxyphenyl)ethyl)-6-methoxy-1,2,3,4-tetrahydroisoquinolin-7-ol.TFA (125)

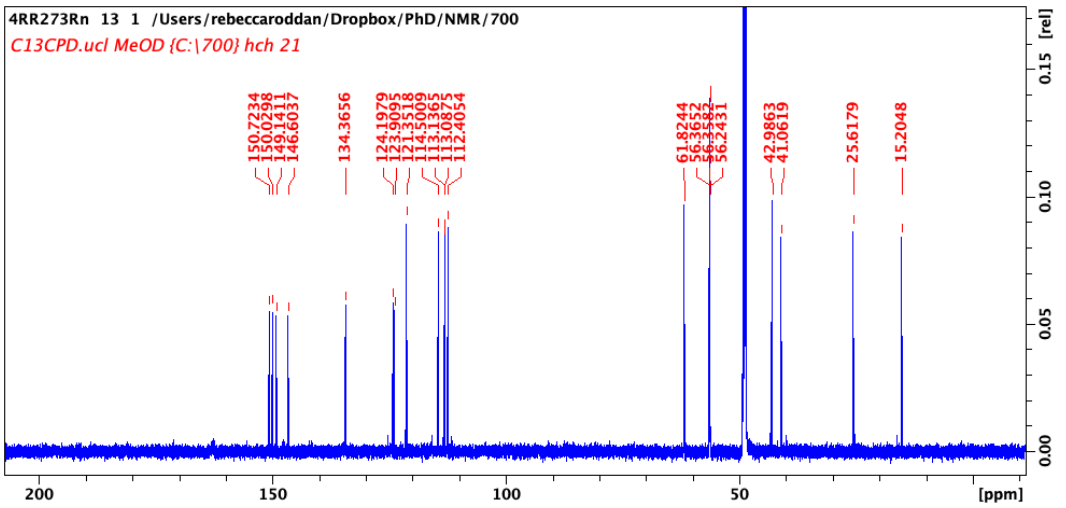


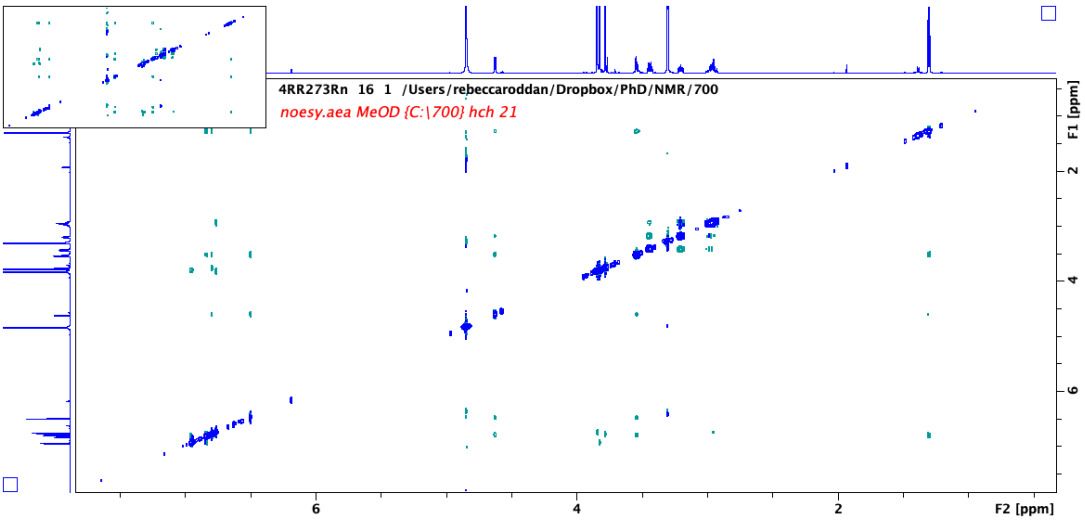




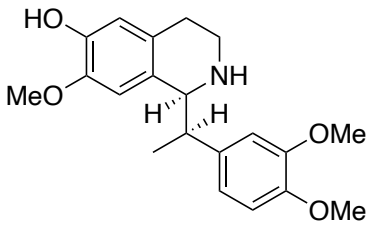

4RR273Rn 10 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
PROTON.ucl MeOD {C:\700} hch 21
1.0157
2.7794
3.2250
2.8794
1.3396
1.0905
4.6
4.4
4.2
4.0
3.8
3.6
[ppm]
[rel]

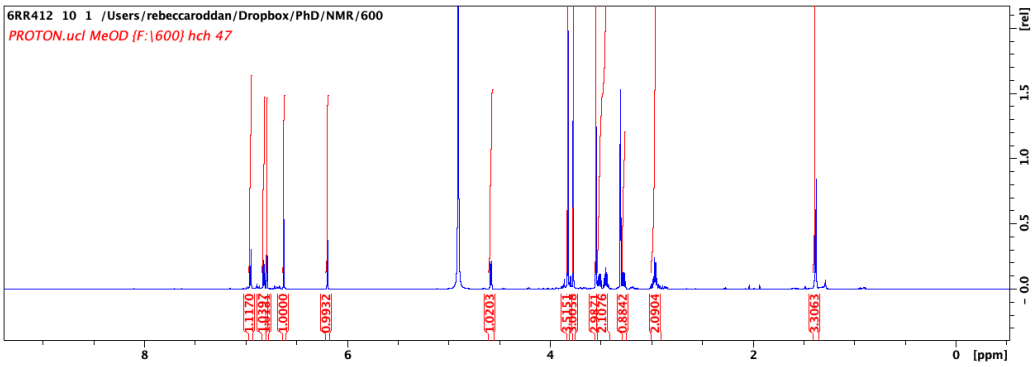

4RR273Rn 10 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
PROTON.ucl MeOD {C:\700} hch 21
1.0152
2.2126
3.0332
3.0
2.5
2.0
1.5
[ppm]
[rel]

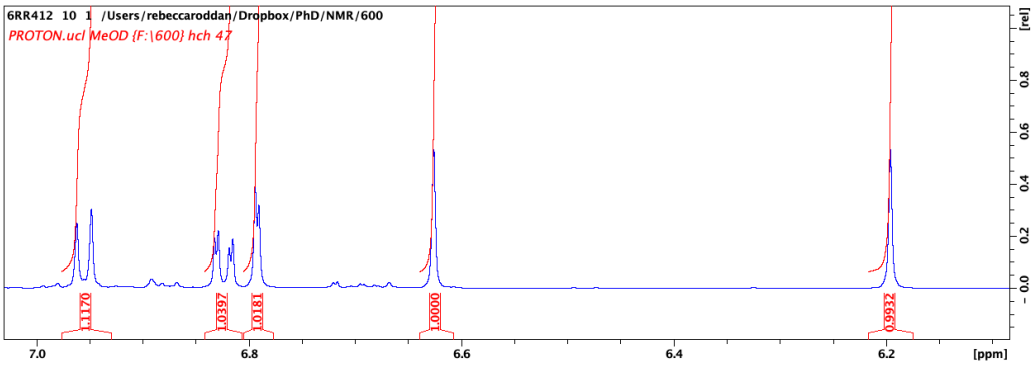

4RR273Rn 13 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
C13CPD.ucl MeOD {C:\700} hch 21
150.7234
150.0298
149.1411
146.6037
134.3656
124.1979
123.9095
121.3518
114.5009
113.1365
113.0875
112.4054
61.8244
56.3632
56.2432
56.2431
42.9863
41.0619
25.6179
15.2048
200
150
100
50
[ppm]
[rel]

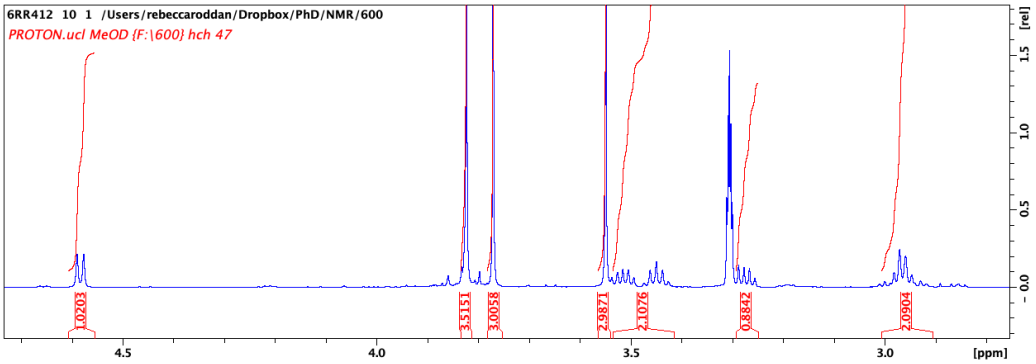

4RR273Rn 16 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
noesy.aea MeOD {C:\700} hch 21
F1 [ppm]
F2 [ppm]
6
4
2

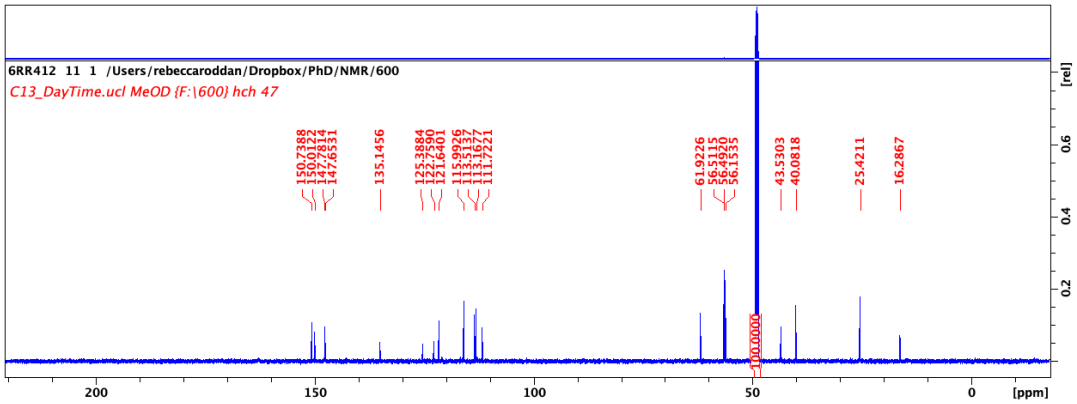
1-((1*S*)-1-((*R*)-1-(3,4-dimethoxyphenyl)ethyl)-6-hydroxy-7-methoxy-1,2,3,4-tetrahydro-2λ^4-isoquinolin-2-yl)-2,2,2-trifluoroethan-1-one.TFA (127)

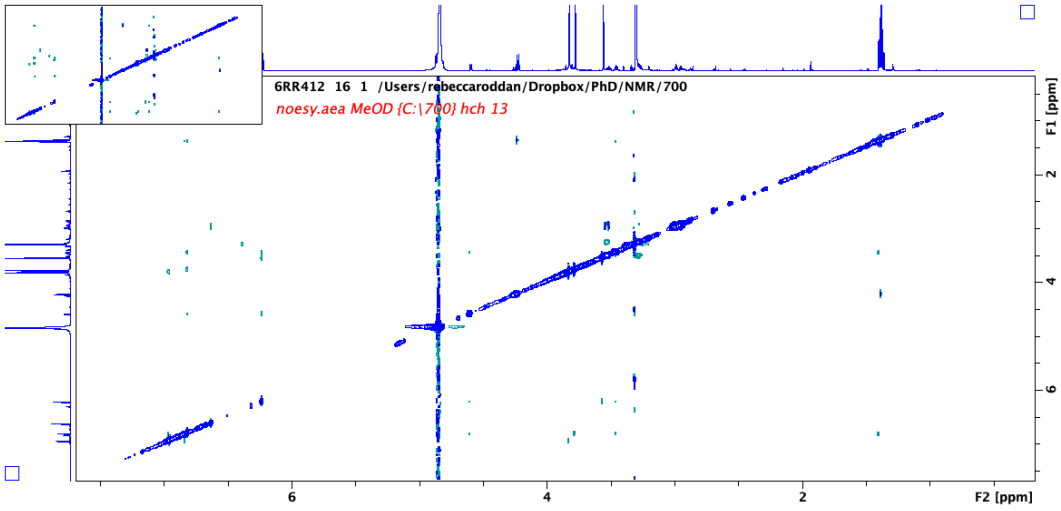




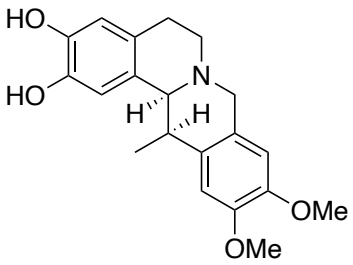


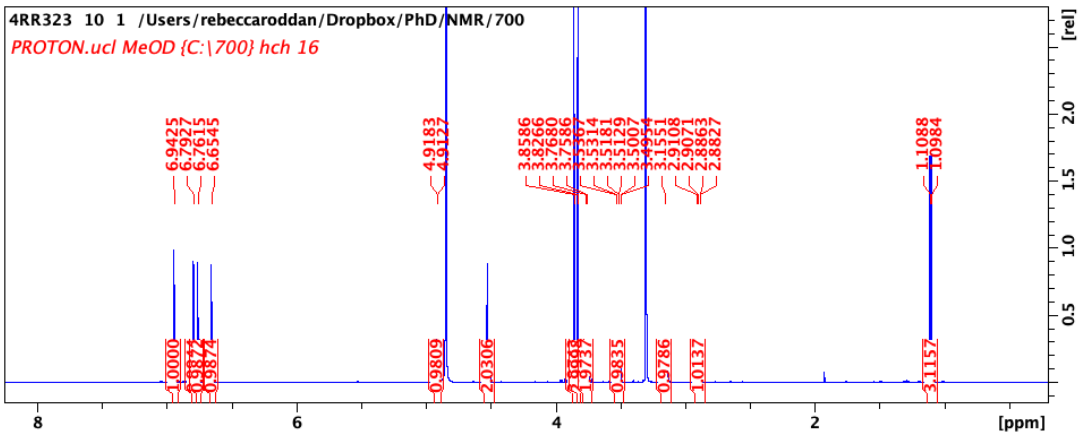


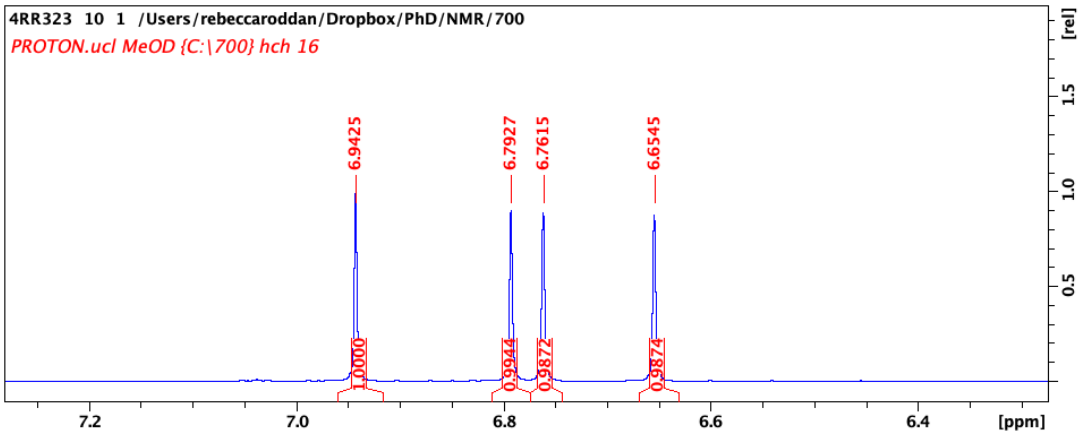

6RR412 11 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/600
C13_DayTime.ucl MeOD {F:\600} hch 47
150.7388
150.0172
147.7814
147.6531
135.1456
125.3884
122.2590
121.6401
115.9926
113.3737
113.1677
111.7221
61.9226
56.5115
56.4920
56.1535
43.5303
40.0818
25.4211
16.2867
100.0000
[rel]
0.2
0.4
0.6
200
150
100
50
0
[ppm]

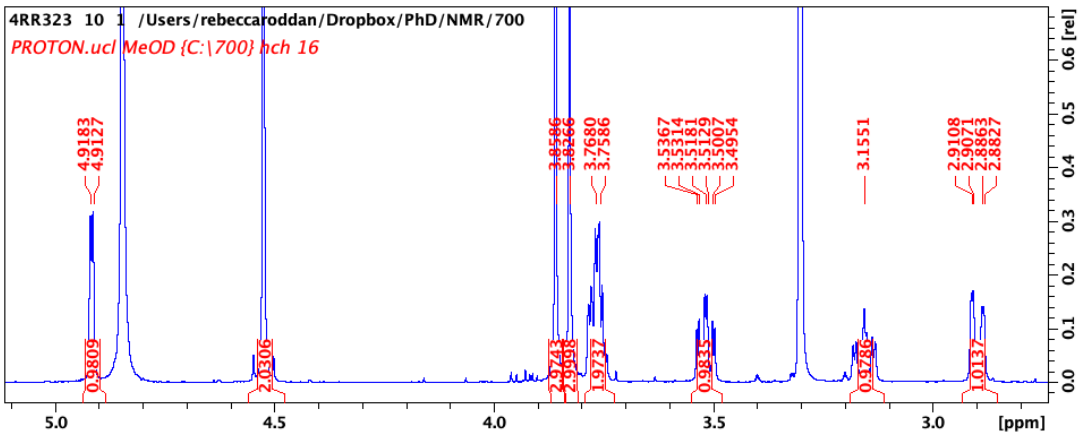

6RR412 16 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
noesy.aea MeOD {C:\700} hch 13
F1 [ppm]
2
4
6
6
4
2
F2 [ppm]

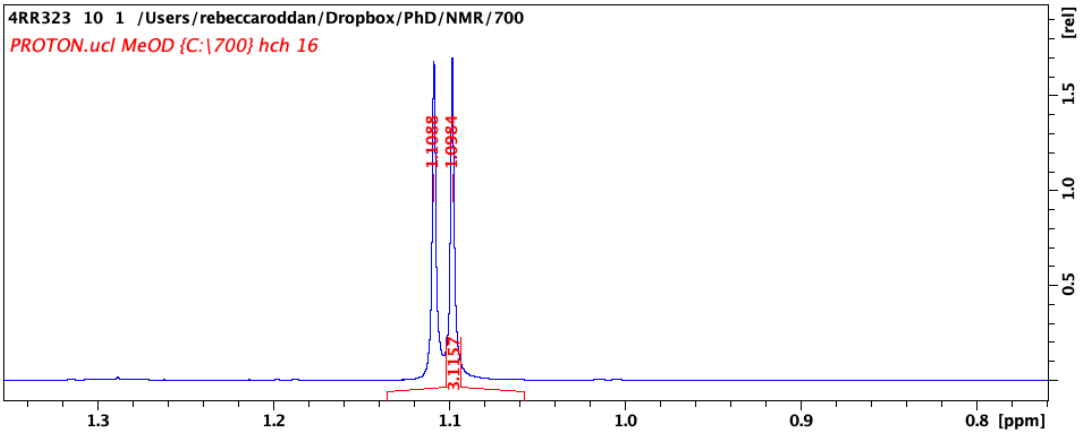
(13*R*,13a*S*)-10,11-Dimethoxy-13-methyl-5,8,13,13a-tetrahydro-6*H*-isoquinolino[3,2-*a*]isoquinoline-2,3-diol.TFA (129)

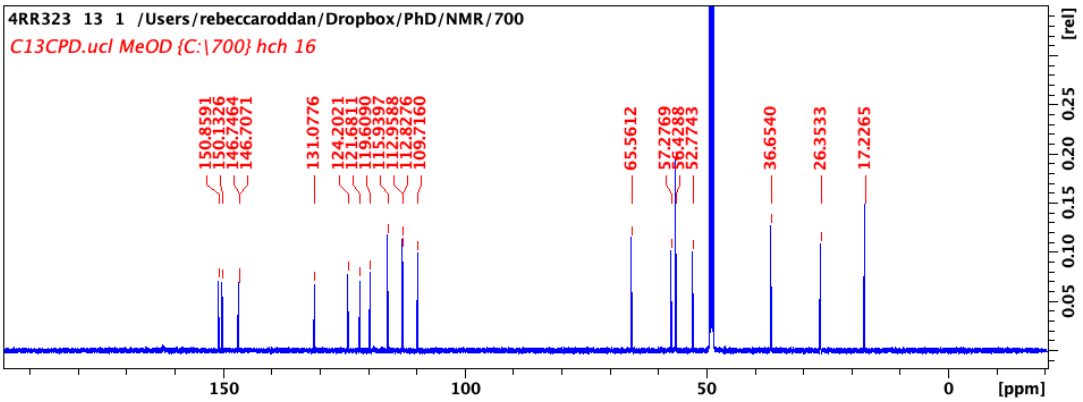


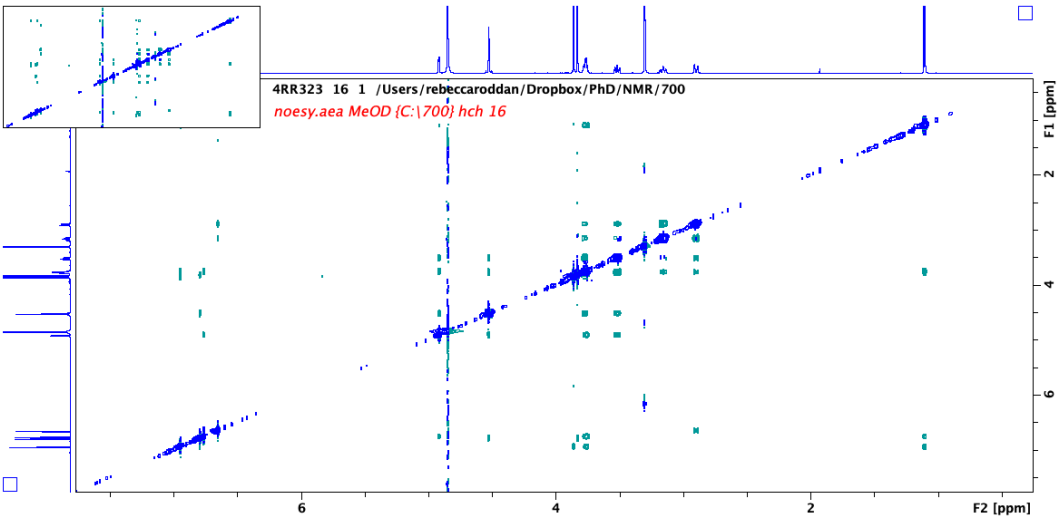


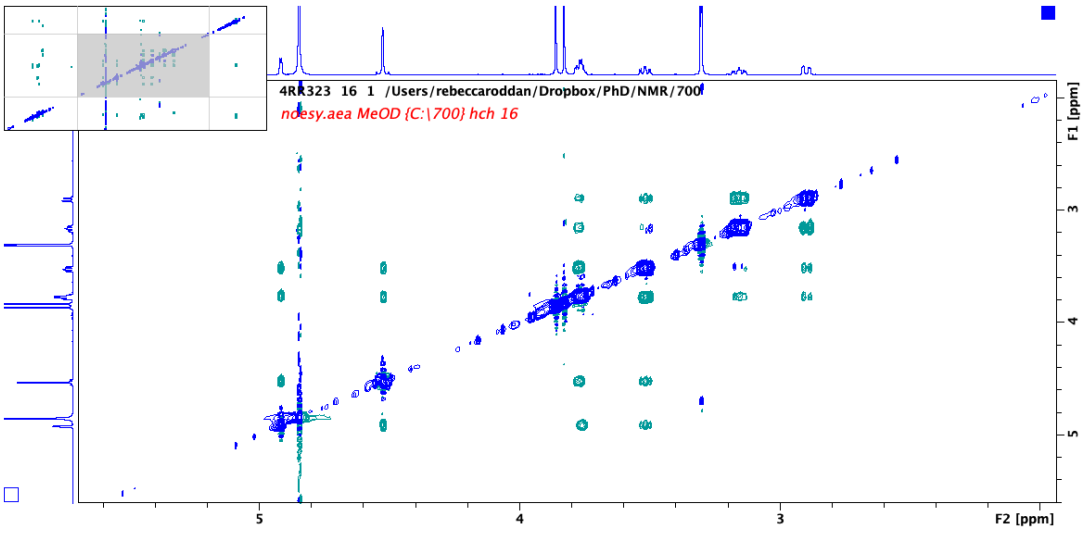




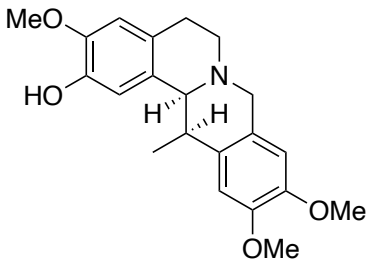

4RR323 10 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
PROTON.ucl MeOD {C:\700} hch 16
1.1088
1.0984
3.1157
[rel]
1.5
1.0
0.5
1.3
1.2
1.1
1.0
0.9
0.8 [ppm]

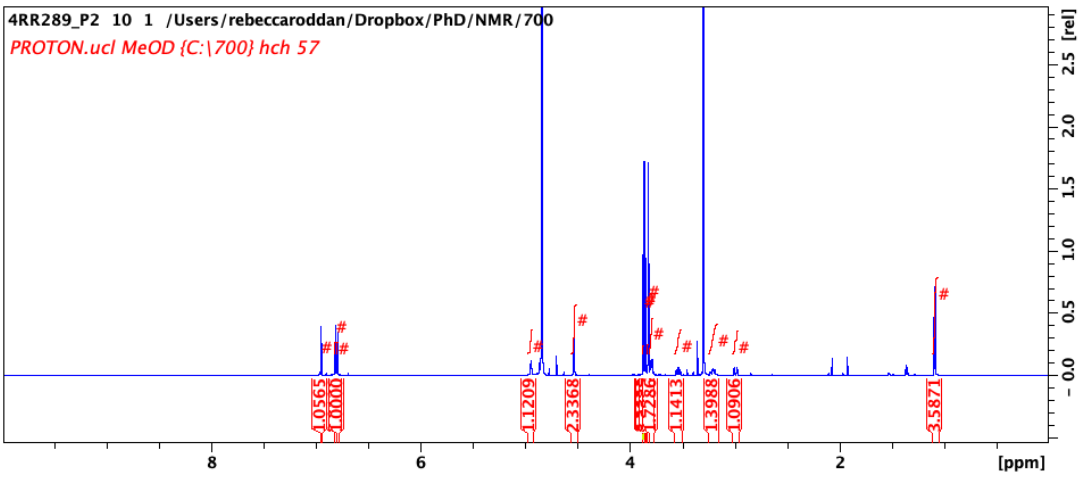

4RR323 13 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
C13CPD.ucl MeOD {C:\700} hch 16
150.8591
150.1326
146.7464
146.7071
131.0776
124.2021
121.6811
119.6090
115.9397
112.9588
112.8276
109.7160
65.5612
57.2769
56.4288
52.7743
36.6540
26.3533
17.2265
[rel]
0.25
0.20
0.15
0.10
0.05
150
100
50
0
[ppm]

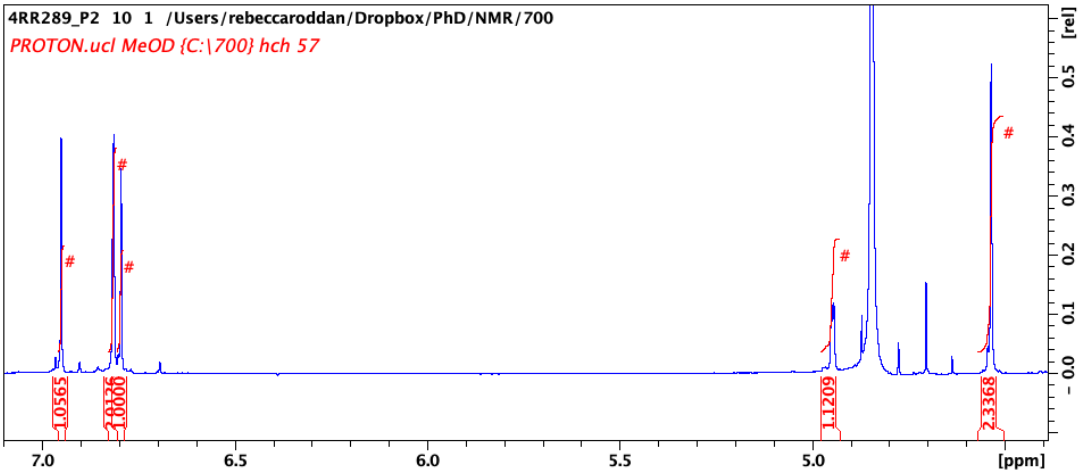

4RR323 16 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
noesy.aea MeOD {C:\700} hch 16
F1 [ppm]
2
4
6
6
4
2
F2 [ppm]

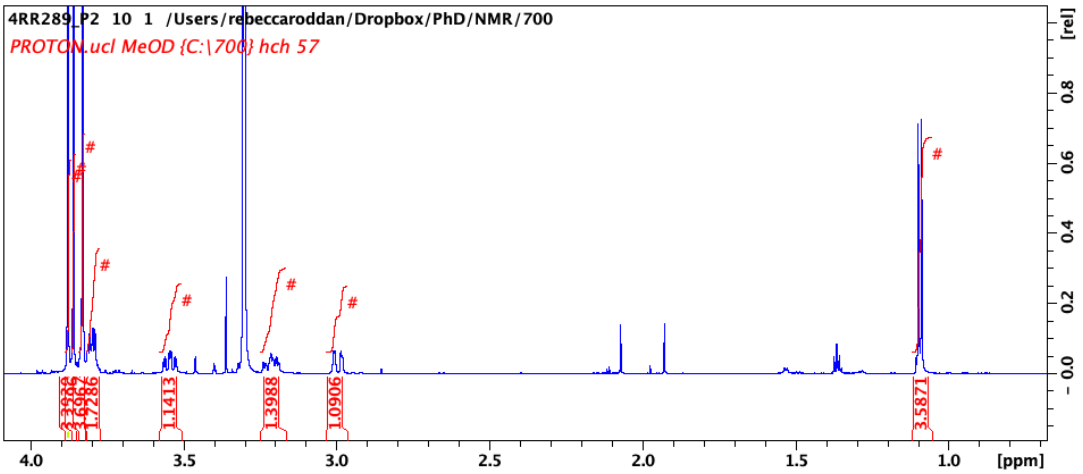

4RR323 16 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
noesy.aea MeOD {C:\700} hch 16
F1 [ppm]
F2 [ppm]
5
4
3

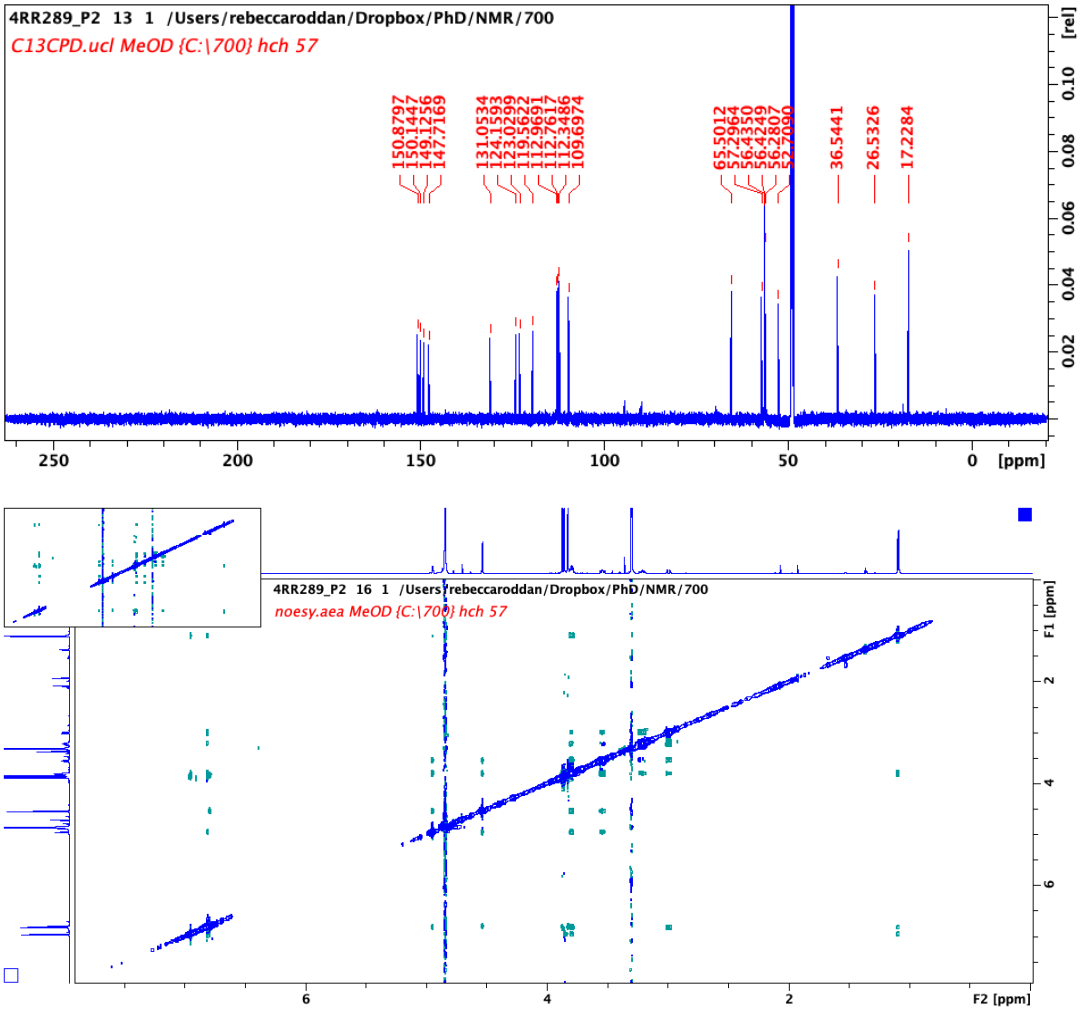
(13*R*,13a*S*)-3,10,11-Trimethoxy-13-methyl-5,8,13,13a-tetrahydro-6*H*-isoquinolino[3,2-*a*]isoquinolin-2-ol.TFA (130)



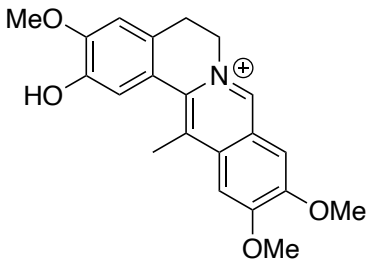


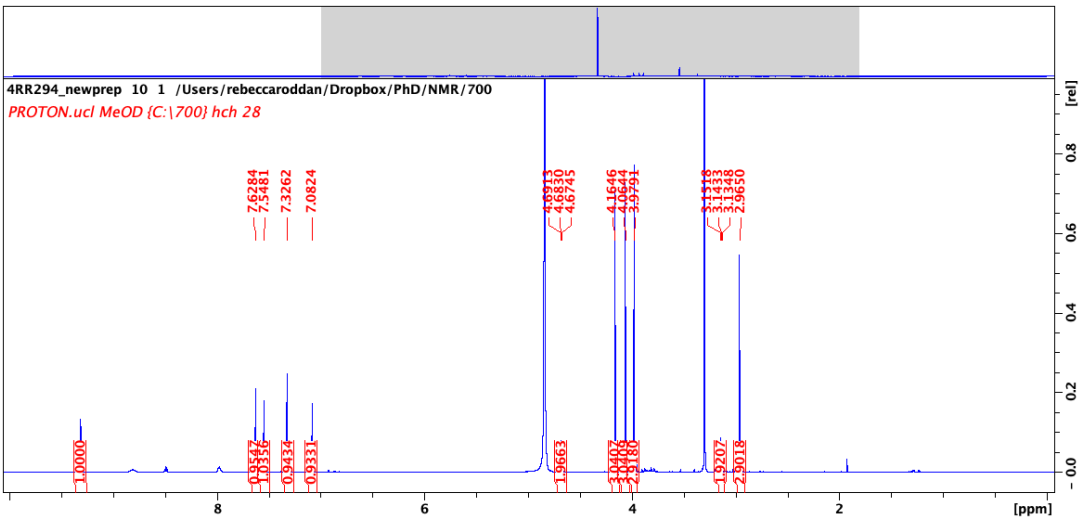


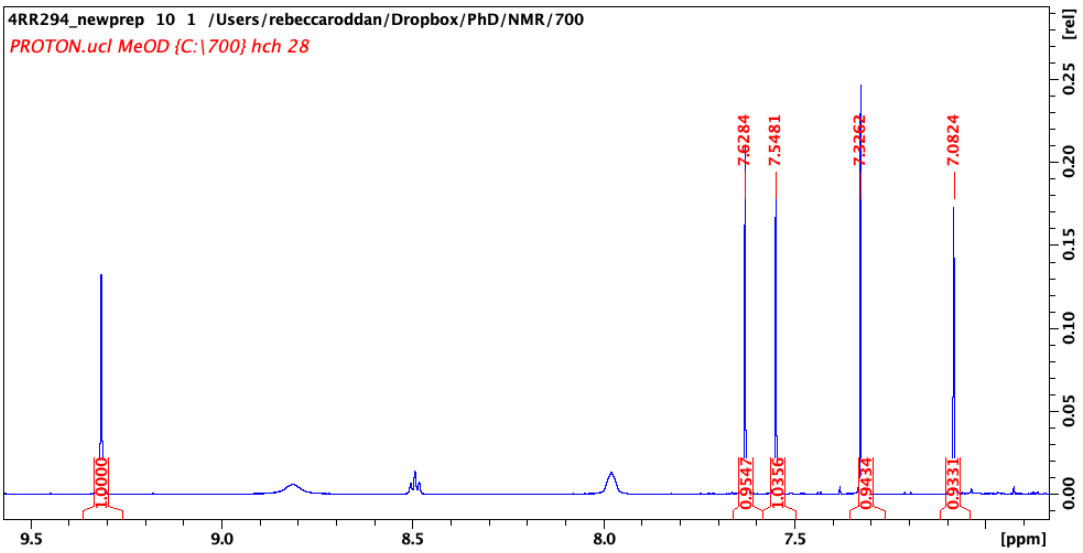


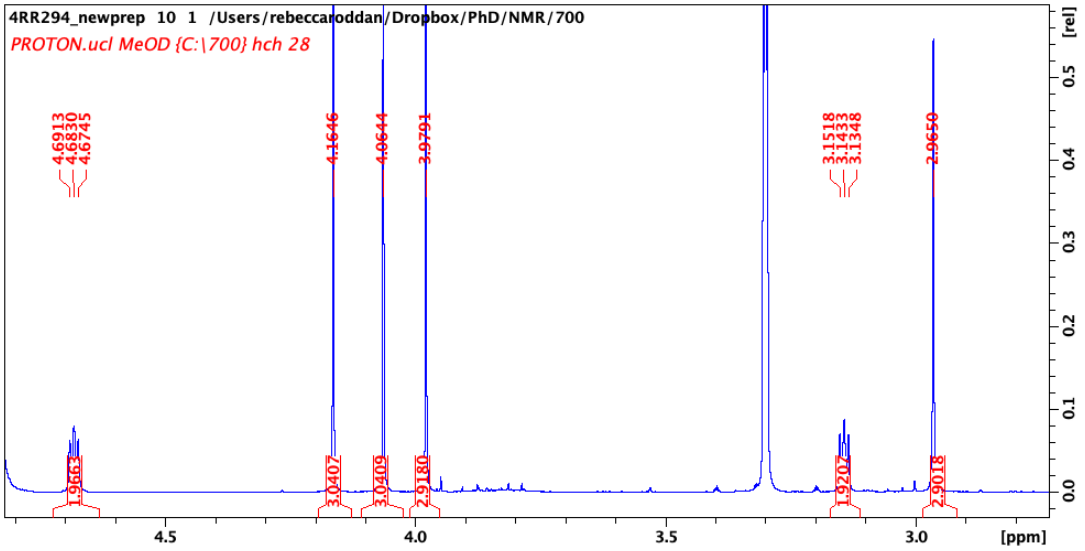

4RR289_P2 13 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
C13CPD.ucl MeOD {C:\700} hch 57
150.8797
150.1447
149.1256
147.7169
131.0534
124.1583
123.0299
119.5622
112.9691
112.7617
112.3486
109.6974
65.5012
57.2964
56.4350
56.4249
56.2807
52.7090
36.5441
26.5326
17.2284
[rel]
0.10
0.08
0.06
0.04
0.02
250
200
150
100
50
0
[ppm]
4RR289_P2 16 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
noesy.aea MeOD {C:\700} hch 57
F1 [ppm]
2
4
6
6
4
2
F2 [ppm]

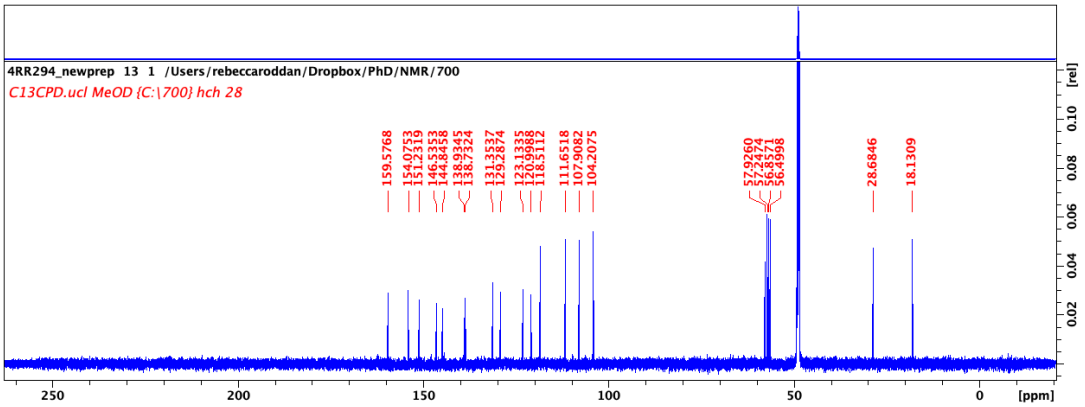
2-Hydroxy-3,10,11-Trimethoxy-13-methyl-5,6-dihydroisoquinolino[3,2-*a*]isoquinolin-7-ium (131)

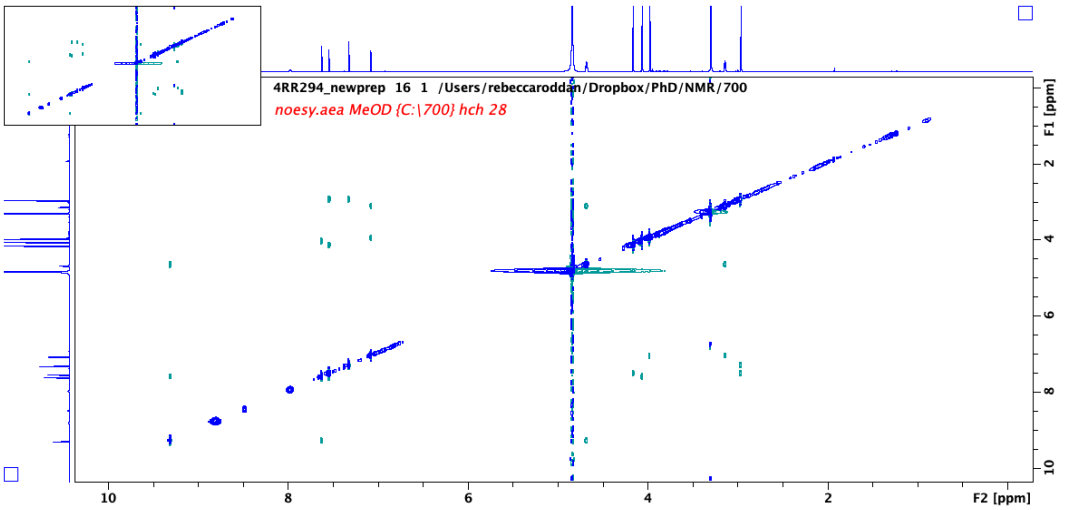




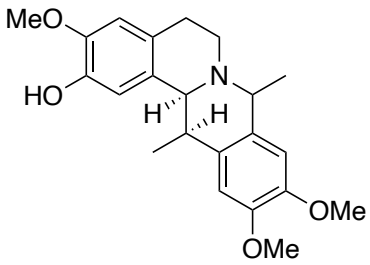


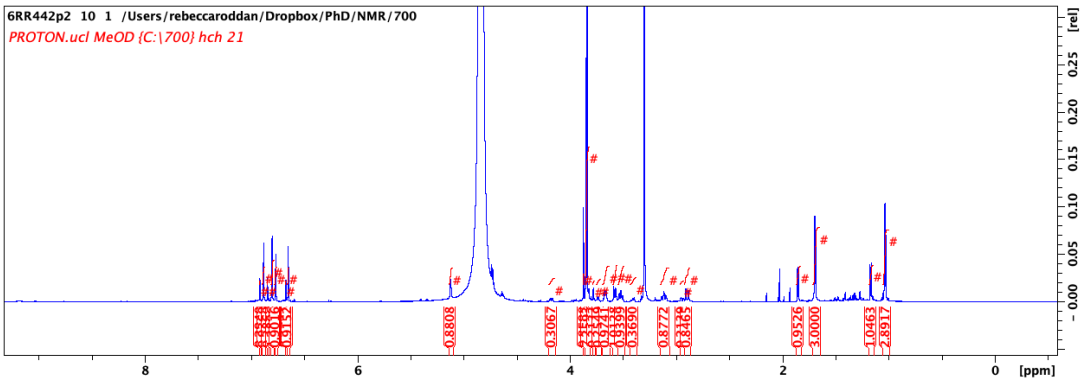

4RR294_newprep 10 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
PROTON.ucl MeOD {C:\700} hch 28
4.6913
4.6830
4.6745
4.1646
4.0644
3.9791
3.1518
3.1433
3.1348
2.9650
1.9663
3.0407
3.0409
2.9180
1.9207
2.9018
[rel]
[ppm]

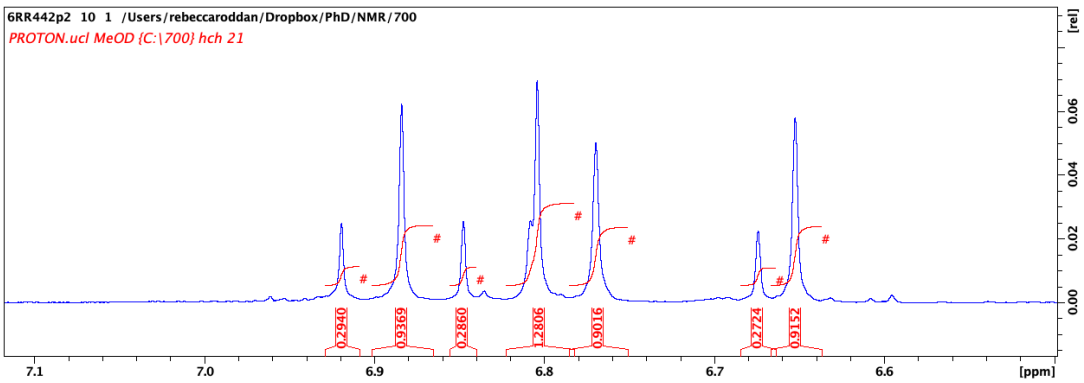

4RR294_newprep 13 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
C13CPD.ucl MeOD {C:\700} hch 28
159.5768
154.0753
151.2319
146.5353
144.8458
138.9345
138.7324
131.3537
129.2874
123.1335
120.9988
118.5112
111.6518
107.9082
104.2075
57.9260
57.2474
56.8571
56.4998
28.6846
18.1309
[rel]
[ppm]

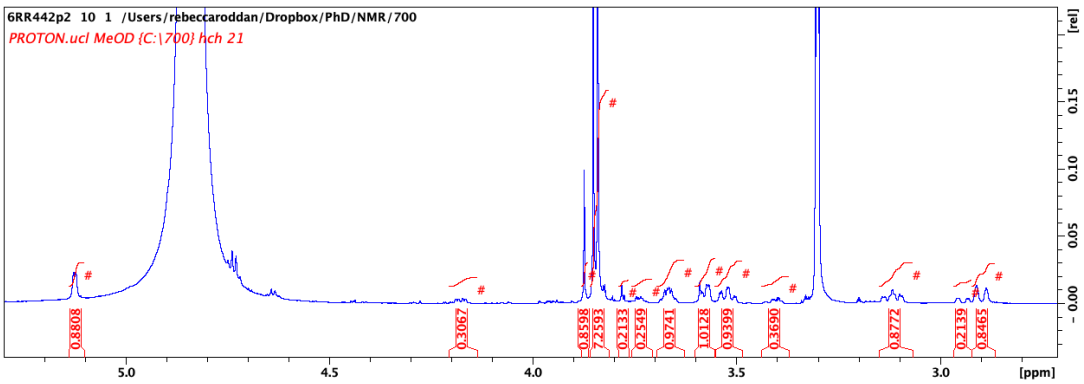

4RR294_newprep 16 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
noesy.aea MeOD {C:\700} hch 28
F1 [ppm]
F2 [ppm]

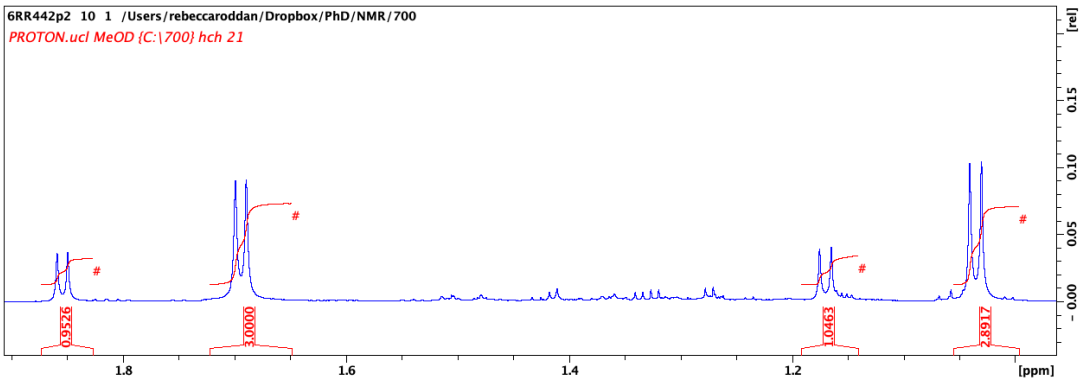
(8*SR*,13*R*,13a*S*)-10,11-Dimethoxy-8,13-dimethyl-5,8,13,13a-tetrahydro-6*H*-isoquinolino[3,2-*a*]isoquinoline-2,3-diol.TFA (140)

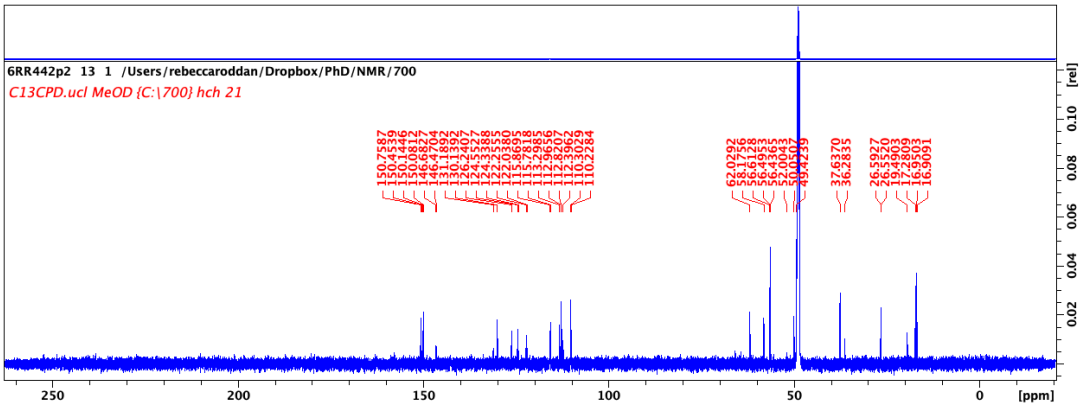


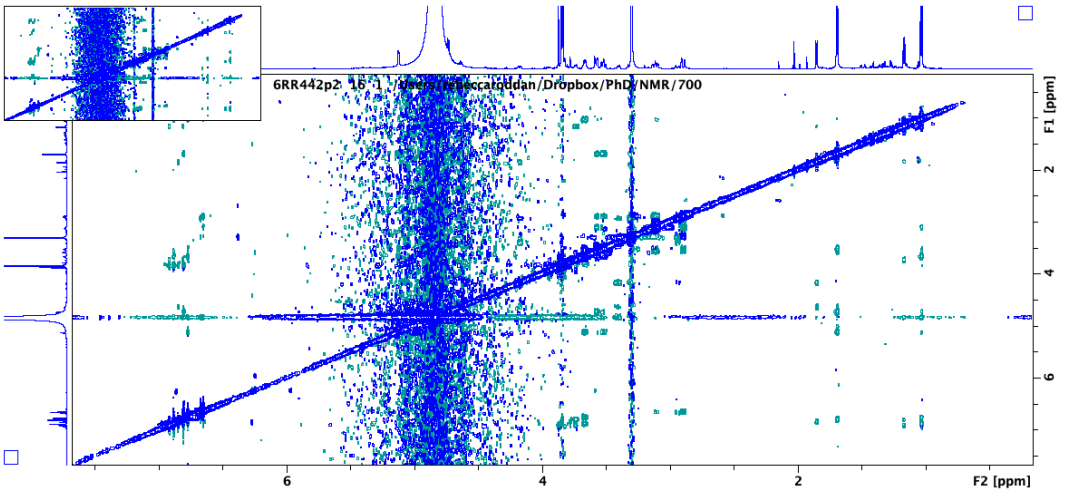


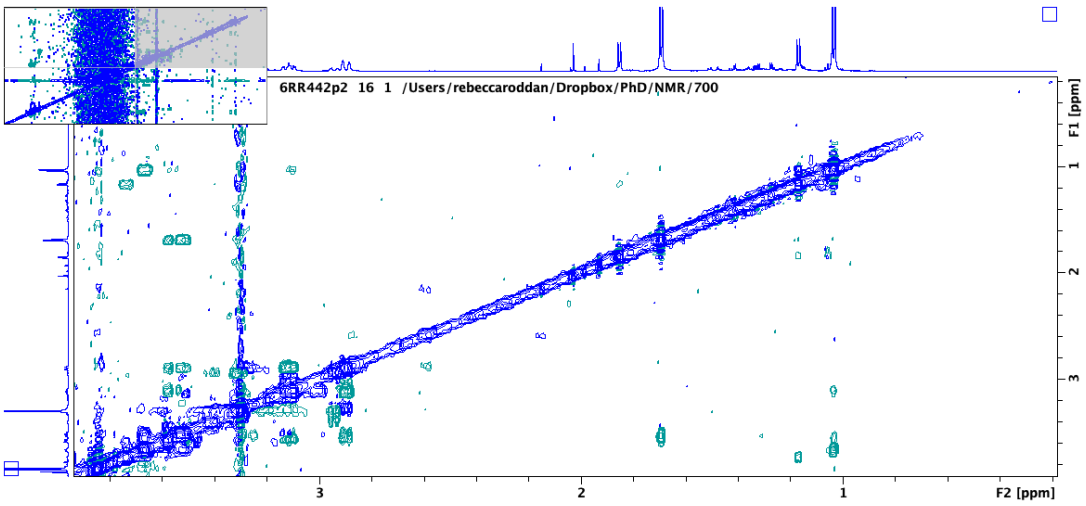





6RR442p2 10 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
PROTON.ucl MeOD {C:\700} hch 21
0.9526
3.0000
1.0463
2.8917
1.8
1.6
1.4
1.2
[ppm]
[rel]


6RR442p2 13 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
C13CPD.ucl MeOD {C:\700} hch 21
250
200
150
100
50
0
[ppm]
[rel]


6RR442p2 16 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
F1 [ppm]
F2 [ppm]
6
4
2


6RR442p2 16 1 /Users/rebeccaroddan/Dropbox/PhD/NMR/700
F1 [ppm]
F2 [ppm]
3
2
1

9.6 Multiple sequence alignment of NCSs

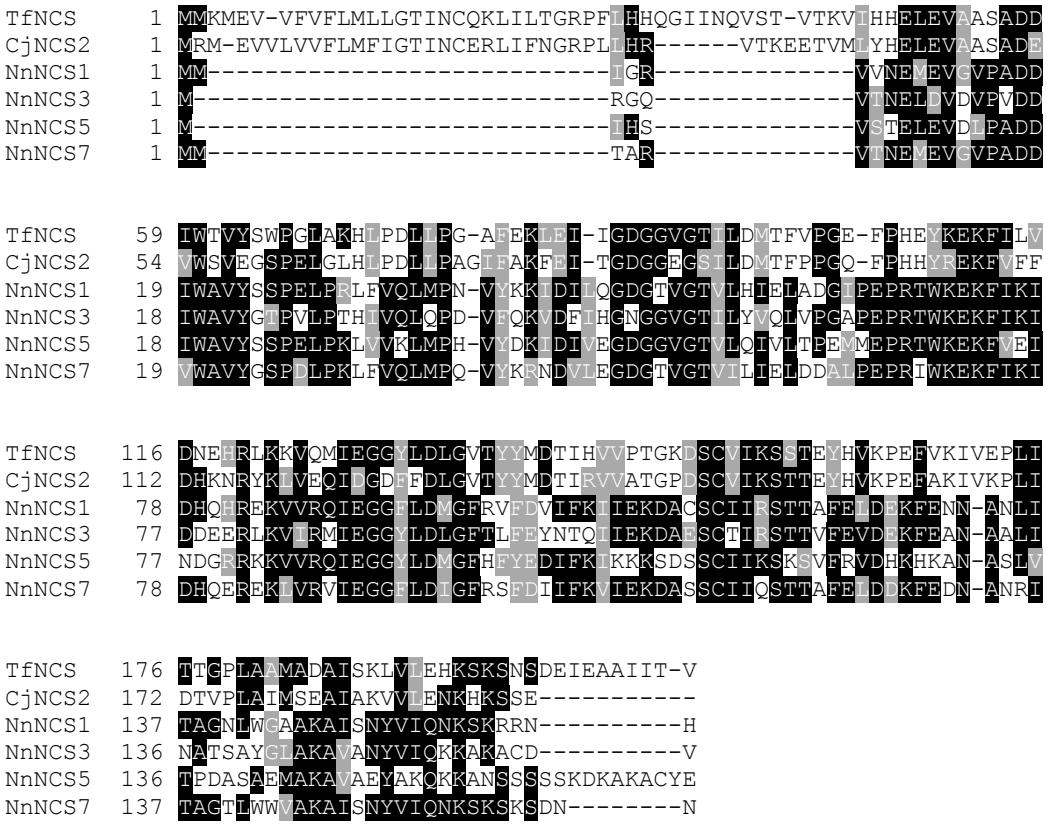


Figure 9.56: Multiple sequence alignment of NCSs. Protein sequences were compared with *Tf*NCS. Black indicates conserved residues and grey indicates residue similarity.

Chapter 10: Bibliography

1 J. M. Hagel and P. J. Facchini, *Plant Cell Physiol.*, 2013, **54**, 647–672.

2 R. E. Loder, *Anaesthesia*, 1969, **24**, 355–358.

3 K. Iwasa, M. Moriyasu, Y. Tachibana, H.-S. Kim, Y. Wataya, W. Wiegrebe, K. F. Bastow, L. M. Cosentino, M. Kozuka and K.-H. Lee, *Bioorganic Med. Chem.*, 2001, **9**, 2871–2884.

4 M. Zhong, Y. Liu, J. Liu, D. Di, M. Xu, Y. Yang, W. Li, Y. Chen and J. Liu, *Molecules*, 2014, **19**, 12099–12115.

5 C. Veeresham, *J. Adv. Pharm. Technol. Res.*, 2012, **3**, 200–1.

6 S. C. Farrow, J. M. Hagel, G. A. W. Beaudoin, D. C. Burns and P. J. Facchini, *Nat. Chem. Biol.*, 2015, **11**, 728–734.

7 T. Winzer, M. Kern, A. J. King, T. R. Larson, R. I. Teodor, S. L. Donninger, Y. Li, A. A. Dowle, J. Cartwright, R. Bates, D. Ashford, J. Thomas, C. Walker, T. A. Bowser and I. A. Graham, *Science (80-.).*, 2015, **349**, 309–312.

8 A. Pictet and T. Spengler, *Ber. Dtsch. Chem. Ges.*, 1911, **43**, 2030–2036.

9 E. Winterstein and G. Trier, *Die Alkaloide. Eine Monographie der naturlichen Basen*, Borntraeger, Berlin, Germany, 1910.

10 M. L. Wilson and C. J. Coscia, *J. Am. Chem. Soc.*, 1975, **97**, 431–432.

11 A. R. Battersby, R. C. F. Jones and R. Kazlauskas, *Tetrahedron Lett.*, 1975, **16**, 1873–1876.

12 M. Rueffer, H. El-Shagi, N. Nagakura and M. H. Zenk, *FEBS Lett.* , 1981, **129**, 5–9.

13 H.-M. Schumacher, M. Ruffer, N. Nagakura and M. H. Zenk, *Planta*, 1983, **48**, 212–220.

14 R. Stadler, T. M. Kutchan and M. H. Zenk, *Phytochemistry*, 1989, **28**, 1083–1086.

15 T. Pesnot, M. C. Gershater, J. M. Ward and H. C. Hailes, *Chem. Commun.*, 2011, **47**, 3242–3244.

16 N. Samanani and P. J. Facchini, *Planta*, 2001, **213**, 898–906.

17 N. Samanani, D. K. Liscombe and P. J. Facchini, *Plant J.*, 2004, **40**, 302–313.

18 T. M. Kutchan, M. Rush, C. J. Coscia and E. A. Doisy, *Plant Physiol*, 1986, **81**, 161–166.

19 R. Roddan, J. M. Ward, N. H. Keep and H. C. Hailes, *Curr. Opin. Chem. Biol.*, 2020, **55**, 69–76.

20 H. Lechner, P. Soriano, R. Poschner, H. C. Hailes, J. M. Ward and W. Kroutil, *Biotechnol. J.*, 2018, **13**, 1–9.

21 B. M. Ruff, S. Bräse and S. E. O'Connor, *Tetrahedron Lett.*, 2012, **53**, 1071–1074.

22 T. Pesnot, M. C. Gershater, J. M. Ward and H. C. Hailes, *Adv. Synth. Catal.*, 2012, **354**, 2997–3008.

23 N. Samanani and P. J. Facchini, *J. Biol. Chem.*, 2002, **277**, 22878–33883.

24 L. Y. P. Luk, S. Bunn, D. K. Liscombe, P. J. Facchini and M. E. Tanner, *Biochemistry*, 2007, **46**, 10153–10161.

25 H. Berkner, J. Engelhorn, D. K. Liscombe, K. Schweimer, B. M. Wöhrl, P. J. Facchini, P. Rösch and I. Matečko, *Protein Expr. Purif.*, 2007, **56**, 197–204.

26 H. Berkner, K. Schweimer, I. Matecko and P. Osch, *Biochem. J*, 2008, **413**, 281–290.

27 A. Ilari, S. Franceschini, A. Bonamore, F. Arenghi, B. Botta, A. Macone, A. Pasquo, L. Bellucci and A. Boffi, *J. Biol. Chem.*, 2009, **284**, 897–904.

28 B. R. Lichman, M. C. Gershater, E. D. Lamming, T. Pesnot, A. Sula, N. H. Keep, H. C. Hailes, J. M. Ward and C. J. M. Ward, *FEBS J.*, 2015, **282**, 1137–1151.

29 A. Mozzarelli and G. L. Rossi, *Annu. Rev. Biophys. Biomol. Struct.*, 1996, **25**, 343–365.

30 B. R. Lichman, A. Sula, T. Pesnot, H. C. Hailes, J. M. Ward and N. H. Keep, *Biochemistry*, 2017, **56**, 5274–5277.

31 B. R. Lichman, University College London, 2015.

32 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E.

C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.

33 X. Sheng and F. Himo, *J. Am. Chem. Soc.*, 2019, **141**, 11230–11238.

34 A. Bar-Even, E. Noor, Y. Savir, W. Liebermeister, D. Davidi, D. S. Tawfik and R. Milo, *Biochemistry*, 2011, **50**, 4402–4410.

35 M. Nishihachijo, Y. Hirai, S. Kawano, A. Nishiyama, H. Minami, T. Katayama, Y. Yasohara, F. Sato and H. Kumagai, *Biosci. Biotechnol. Biochem.*, 2014, **78**, 701–707.

36 B. R. Lichman, J. Zhao, H. C. Hailes and J. M. Ward, *Nat. Commun.*, 2017, **8**, 14883.

37 J. Zhao, D. Mendez-Sanchez, R. Roddan, J. M. Ward and H. C. Hailes, *ACS Catal.*, 2021, **11**, 131–138.

38 A. I. Meyers, M. A. Gonzalez, V. Struzka, A. Akahane, J. Guiles and J. S. Warmus, *Tetrahedron Lett.*, 1991, **32**, 5501–5504.

39 H.-T. Luu, S. Wiesler, G. Frey and J. Streuff, *Org. Lett.*, 2015, **17**, 2478–2481.

40 H. Fernandes, K. Michalska, M. Sikorski and M. Jaskolski, *FEBS J.*, 2013, **280**, 1169–1199.

41 C. Radauer, P. Lackner and H. Breiteneder, *BMC Evol. Biol.*, 2008, **8**, 286–294.

42 B. D. Ames, T. P. Korman, W. Zhang, P. Smith, T. Vu, Y. Tang and S.-C. Tsai, *PNAS*, 2008, **105**, 5349–5354.

43 J. S. Morris, K. M. Caldo, S. Liang and P. J. Facchini, *ChemBioChem*, 2021, **22**, 264–287.

44 A. Singh, M. Massicotte, A. Garand, L. Tousignant, V. Ouellette, G. Bérubé and I. Desgagné-penix, *BMC Plant Biol.*, 2018, **18**, 1–12.

45 D. K. Liscombe, B. P. MacLeod, N. Loukanina, O. I. Nandi and P. J. Facchini, *Phytochemistry*, 2005, **66**, 1374–1393.

46 S. Vimolmangkang, X. Deng, A. Owiti, T. Meelaph, C. Ogutu and Y. Han, *Sci. Rep.*, 2016, **6**, 26323.

47 B. R. Sharma, L. N. S. Gautam, D. Adhikari and R. Karki, *Phyther. Res.*, 2016,

26, 3–26.

48 H. Hong, Y.-I. Lee and D. Jin, *Microchem. J.*, 2010, **96**, 374–379.

49 S. Nishibe, H. Tsukamoto, H. Kinoshita, S. Kitagawa and A. Sakushima, *J. Nat. Prod.*, 1986, **49**, 547–548.

50 Y. Kashiwada, A. Aoshima, Y. Ikeshiro, Y.-P. Chen, H. Furukawa, M. Itoigawa, T. Fujioka, K. Mihashi, L. M. Cosentino, S. L. Morris-Natschke and K.-H. Lee, 2004, **13**, 443–448.

51 H. Minami, E. Dubouzet, K. Iwasa and F. Sato, *J. Biol. Chem.*, 2007, **282**, 6274–6282.

52 J. D. Guzman, T. Pesnot, D. A. Barrera, H. M. Davies, E. Mcmahon, D. Evangelopoulos, P. N. Mortazavi, T. Munshi, A. Maitra, E. D. Lamming, R. Angell, M. C. Gershater, J. M. Redmond, D. Needham, J. M. Ward, L. E. Cuca, H. C. Hailes and S. Bhakta, *J. Antimicrob. Chemother.*, 2015, **70**, 1691–1703.

53 H. Hatano, F. Takekawa, K. Hashimoto, M. Ishihara, M. Kawase, C. Qing, W. Qin-Tao and H. Sakagami, *Anticancer Res.*, 2009, **29**, 3079–3086.

54 J. Jankovic and J. Beach, *Neurology*, 1997, **48**, 358–362.

55 K. Ikeda, S. Kobayashi, M. Suzuki, K. Miyata, M. Takeuchi, T. Yamada and K. Honda, *Naunyn. Schmiedebergs. Arch. Pharmacol.*, 2002, **366**, 97–103.

56 J. McConathy, C. B. Nemeroff and M. J. Owens, *Essent. Psychopharmacol.* , 2004, **5**, 297–306.

57 M. K. Pyo, D.-H. Lee, D.-H. Kim, J.-H. Lee, J.-C. Moon, K. C. Chang and H. S. Yun-Choi, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 4110–4114.

58 R. D. Parra and J. Maresh, *Comput. Theor. Chem.*, 2016, **1082**, 1–10.

59 J. Zhao, D. Méndez-Sánchez, J. M. Ward and H. C. Hailes, *J. Org. Chem.*, 2019, **84**, 7702–7710.

60 I. Almodovar, M. C. Rezende, B. K. Cassels and M. García-Arriagada, *J. Phys. Org. Chem.*, 2017, **30**, 1–6.

61 V. Erdmann, B. R. Lichman, J. Zhao, R. C. Simon, W. Kroutil, J. M. Ward, H. C.

Hailes and D. Rother, *Angew. Chemie*, 2017, **129**, 12677–12681.

62 A. Bischler and B. Napieralkski, *Ber.*, 1893, **26**, 1903–1908.

63 J. Seayad, A. M. Seayad and B. List, *J. Am. Chem. Soc.*, 2006, **128**, 1086–1087.

64 E. Mons, M. J. Wanner, S. Ingemann, J. H. Van Maarseveen and H. Hiemstra, *J. Org. Chem.* , 2014, **79**, 7380–7390.

65 D. L. Comins, P. M. Thakker, M. F. Baevsky and M. M. Badawi, *Tetrahedron*, 1997, **53**, 16327–16340.

66 M. M Heravi, V. Zadsirjan and M. Malmir, *Molecules*, 2018, **23**, 1–48.

67 D. Ghislieri, A. P. Green, M. Pontini, S. C. Willies, I. Rowles, A. Frank, G. Grogan and N. J. Turner, *J. Am. Chem. Soc.*, 2013, **135**, 10863–10869.

68 F. Leipold, S. Hussain, D. Ghislieri and N. J. Turner, *ChemCatChem*, 2013, **5**, 3505–3508.

69 S. Hussain, F. Leipold, H. Man, E. Wells, S. P. France, K. R. Mulholland, G. Grogan and N. J. Turner, *ChemCatChem*, 2015, **7**, 579–583.

70 M. Chang, W. Li, X. Zhang, M. Chang, W. Li and X. Zhang, *Angew. Chem. Int. Ed*, 2011, **50**, 10679–10681.

71 L. Yang, J. Zhu, C. Sun, Z. Deng and X. Qu, *Chem. Sci.*, 2020, **11**, 364–371.

72 S. C. Farrow, J. M. Hagel, G. A. W. Beaudoin, D. C. Burns and P. J. Facchini, *Nat. Chem. Biol.*, 2015, **11**, 728–734.

73 S. Galanie, K. Thodey, I. J. Trenchard, M. F. Interrante and C. D. Smolke, *Science (80-.).*, 2015, **349**, 1095–1100.

74 A. Winkler, A. Lyskowski, S. Riedl, M. Puhl, T. M. Kutchan, P. Macheroux and K. Gruber, *Nat. Chem. Biol.*, 2008, **4**, 739–741.

75 T. M. Kutchan and H. Dittrich, *J. Biol. Chem.*, 1995, **270**, 24475–24481.

76 J. H. Schrittwieser, V. Resch, S. Wallner, W. D. Lienhart, J. H. Sattler, J. Resch, P. MacHeroux and W. Kroutil, *J. Org. Chem.*, 2011, **76**, 6703–6714.

77 V. Resch, H. Lechner, J. H. Schrittwieser, S. Wallner, K. Gruber, P. MacHeroux and W. Kroutil, *Chem. - A Eur. J.*, 2012, **18**, 13173–13179.

78 S. Kishimoto, M. Sato, Y. Tsunematsu and K. Watanabe, *Molecules*, 2016, **21**, 1078–1096.

79 S. E. O'Connor and J. J. Maresh, *Nat. Prod. Rep.*, 2006, **23**, 532–547.

80 S. Brown, M. Clastre, V. Courdavault and S. E. O'Connor, *PNAS*, 2015, **112**, 3205–3210.

81 X. Ma, S. Panjikar, J. Koepke, E. Loris and J. Stöckigt, *Plant Cell Online*, 2006, **18**, 907–920.

82 J. J. Maresh, L.-A. Giddings, A. Friedrich, E. A. Loris, S. Panjikar, B. L. Trout, J. Stöckigt, B. Peters and S. E. O' Connor, *J. Am. Chem. Soc.*, 2008, **130**, 710–723.

83 E. Fischereder, D. Pressnitz, W. Kroutil and S. Lutz, *Bioorganic Med. Chem.*, 2014, **22**, 5633–5637.

84 F. Wu, H. Zhu, L. Sun, C. Rajendran, M. Wang, X. Ren, S. Panjikar, A. Cherkasov, H. Zou and J. Stö, *J. Am. Chem. Soc.*, 2011, **134**, 1498–1500.

85 S. Chen, M. C. Galan, C. Coltharp and S. E. O'Connor, *Chem. Biol.*, 2006, **13**, 1137–1141.

86 E. A. Loris, S. Panjikar, M. Ruppert, L. Barleben, M. Unger, H. Schübel and J. Stöckigt, *Chem. Biol.*, 2007, **14**, 979–985.

87 E. McCoy, M. C. Galan and S. E. O'Connor, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 2475–2478.

88 H. Zhu, P. Kerčmar, F. Wu, C. Rajendran, L. Sun, M. Wang and J. Stöckigt, *Curr. Med. Chem.*, 2015, **22**, 1880–1888.

89 D. Pressnitz, E. Fischereder, J. Pletz, C. Kofler, L. Hammerer, K. Hiebler, H. Lechner, N. Richter, E. Eger and W. Kroutil, *Angew. Chem. Int. Ed.*, 2018, **57**, 1–6.

90 E. Eger, A. Simon, M. Sharma, S. Yang, W. B. Breukelaar, G. Grogan, K. N. Houk and W. Kroutil, *J. Am. Chem. Soc.*, 2020, **142**, 792–800.

91 Y. Cai, H. Zhu, Z. Alperstein, W. Yu, A. Cherkasov and H. Zou, *ACS Chem. Biol.*,

2017, **12**, 3086–3092.

92 W. De-Eknamkul, N. Suttipanta and T. M. Kutchan, *Phytochemistry*, 2000, **55**, 177–181.

93 A. K. Evans and C. A. Lowry, *CNS Drug Rev.*, 2007, **13**, 475–501.

94 T. Mori, S. Hoshino, S. Sahashi, T. Matsui, H. Morita and I. Abe, *Chem. Biol.*, 2015, **22**, 898–906.

95 Q. Chen, S. Zhang and Y. Xie, *J. Biotechnol.*, 2018, **281**, 137–143.

96 J. Brevet, D. Borowski and J. Tempe, *Mol. Plant-Microbe Interact.*, 1988, **1**, 75–70.

97 K. Suzuki, N. Tanaka, H. Kamada and I. Yamashita, *Gene*, 2001, **263**, 49–58.

98 V. Kovacova, J. Zluvova, B. Janousek, M. Talianova and B. Vyskot, *PLoS One*, 2014, **9**, 1–23.

99 X. Wang, D. Kong, T. Huang, Z. Deng and S. Lin, *Org. Biomol. Chem*, 2018, **16**, 9124–9128.

100 K. Koketsu, K. Watanabe, H. Suda, H. Oguri and H. Oikawa, *Nat. Chem. Biol.*, 2010, **6**, 408–410.

101 W. Yan, H. M. Ge, G. Wang, N. Jiang, Y. N. Mei, R. Jiang, S. J. Li, C. J. Chen, R. H. Jiao, Q. Xu, S. W. Ng and R. X. Tan, *PNAS*, 2014, **111**, 18138–18143.

102 M.-S. García-Ayllón, D. H. Small, J. Avila, J. Sáez-Valero and K. Tsim, *Front. Mol. Neurosci.*, 2011, **4**, 1–9.

103 M. Kurnik-Łucka, P. Panula, A. Bugajski and K. Gil, *Neurotox Res*, 2018, **33**, 485–514.

104 Y. Deng, W. Maruyama, P. Dostert, T. Takahashi, M. Kawai and M. Naoi E', *J. Chromatogr. B*, 1995, **670**, 47–54.

105 M. Naoi, W. Maruyama, P. Dostert, K. Kohda and T. Kaiya, *Neurosci. Lett.*, 1996, **212**, 183–186.

106 X. Chen, A. Arshad, H. Qing, R. Wang, J. Lu and Y. Deng, *Biologia (Bratisl).*, 2011, **66**, 1183–1188.

107 X. Chen, X. Zheng, S. Ali, M. Guo, R. Zhong, Z. Chen, Y. Zhang, H. Qing and Y. Deng, *ACS Chem. Neurosci.*, 2018, **9**, 1388–1398.

108 S. Arias, T. Terrazas and K. Cameron, *Syst. Bot.*, 2003, **28**, 547–557.

109 C. Djerassi, G. W. Krakower, A. J. Lemin, H. L. Liang, J. S. Mills and R. Villotti, 1958, **80**, 6284–6292.

110 D. G. O'Donovan and H. Horan, *J. Chem. Soc. C Org.*, 1968, **22**, 2791–2795.

111 D. G. O'Donovan and E. Barry, *J. Chem. Sci. Perkin I*, 1974, 2528–2529.

112 D. S. Seigler, *Plant Secondary Metabolism*, Kluwer Academic Publishers, Norwell, Massachusetts, 2nd edn., 1998.

113 T. A. Geissman and D. H. G. Crout, *Organic Chemistry of Secondary Plant Metabolism*, Freeman, Cooper & Company, San Francisco, CA, U.S.A., 1969.

114 K. Leander, B. Luning and E. Ruusa, *Acta Chem. Scand.*, 1969, **23**, 244–248.

115 S. Agurell, I. Granelli, K. Leander, B. Luning and J. Rosenblom, *Acta Chem. Scand. B*, 1974, **28**, 239–243.

116 A. Brossi and S. Teitel, *Helv. Chim. Acta*, 1971, **54**, 1564–1571.

117 B. R. Lichman, E. D. Lamming, T. Pesnot, J. M. Smith, H. C. Hailes and J. M. Ward, *Green Chem.*, 2015, **17**, 852–855.

118 J. Zhao, B. R. Lichman, J. M. Ward and H. C. Hailes, *Chem. Commun.*, 2018, **54**, 1323–1326.

119 R. F. Wang, X. W. Yang, C. M. Ma, S. Q. Cai, J. N. Li and Y. Shoyama, *Heterocycles*, 2004, **63**, 1443–1448.

120 Y. Wang, N. Tappertzhofen, D. Méndez-Sánchez, M. Bawn, B. Lyu, J. Ward and H. C. Hailes, *Angew. Chemie Int. Ed.*, 2019, **131**, 1–7.

121 D. F. Taber, T. D. Neubert and A. L. Rheingold, *J. Am. Chem. Soc.*, 2002, **124**, 12416–12417.

122 K. M. Hawkins and C. D. Smolke, *Nat. Chem. Biol.*, 2008, **4**, 564–573.

123 H. Minami, J.-S. Kim, N. Ikezawa, T. Takemura, T. Katayama, H. Kumagai, F. Sato and Y. Yamada, *PNAS*, 2008, **105**, 7393–7398.

124 J.-S. Kim, A. Nakagawa, Y. Yamazaki, E. Matsumura, T. Koyanagi, H. Minami, T. Katayama, F. Sato and H. Kumagai, *Biosci. Biotechnol. Biochem.*, 2013, **77**, 2166–2168.

125 A. Nakagawa, H. Minami, J. S. Kim, T. Koyanagi, T. Katayama, F. Sato and H. Kumagai, *Nat. Commun.*, 2011, **2**, 1–9.

126 K. Thodey, S. Galanie and C. D. Smolke, *Nat. Chem. Biol.*, 2014, **10**, 837–846.

127 I. J. Trenchard, M. S. Siddiqui, K. Thodey and C. D. Smolke, *Metab. Eng.*, 2015, **31**, 74–83.

128 S. Galanie, K. Thodey, I. J. Trenchard, M. Filsinger Interrante and C. D. Smolke, *Science (80-.).*, 2015, **349**, 1095–1100.

129 Y. Li, S. Li, K. Thodey, I. Trenchard, A. Cravens and C. D. Smolke, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E3922–E3931.

130 M. Pyne, K. Kevvai, P. S. Grewal, L. Narcross, B. Choi, L. Bourgeois, J. E. Dueber and V. J. J. Martin, *DOI: 10.1101/863506*, .

131 P. S. Grewal, J. A. Samson, J. J. Baker, B. Choi and J. E. Dueber, *Nat. Chem. Biol.*, 2021, **17**, 96–103.

132 R. A. Sheldon and J. M. Woodley, *Chem. Rev.*, 2018, **118**, 801–838.

133 F. H. Arnold, *Angew. Chemie - Int. Ed.*, 2018, **57**, 4143–4148.

134 S. B. J. Kan, R. D. Lewis, C. K. and F. H. Arnold, *Science (80-.).*, 2016, **354**, 1048–1051.

135 S. B. Jennifer Kan, X. Huang, Y. Gumulya, K. Chen and F. H. Arnold, *Nature*, 2017, **552**, 132–136.

136 P. S. Huang, S. E. Boyken and D. Baker, *Nature*, 2016, **537**, 320–327.

137 P. Muñoz, S. Huenchuguala, I. Paris and J. Segura-Aguilar, *Parkinsons. Dis.*, 2012, **2012**, 1–13.

138 A. Leggio, E. L. Belsito, G. De Luca, M. L. Di Gioia, V. Leotta, E. Romio, C. Siciliano and A. Liguori, *RSC Adv.*, 2016, **6**, 34468–34475.

139 L. A. Nguyen, H. He and C. Pham-Huy, *Int. J. Biomed. Sci.*, 2006, **2**, 85–100.

140 I. M. Menéndez-Perdomo and P. J. Facchini, *Molecules*, 2018, **23**, 2899–2916.

141 H. Koshiyama, H. Ohkuma, H. Kawaguchi, H.-Y. Hsu and Y.-P. Chen, *Chem. Pharm. Bull.*, 1970, **18**, 2564–2568.

142 V. M. Rekdal, E. N. Bess, J. E. Bisanz, P. J. Turnbaugh and E. P. Balskus, *Science (80-.).*, 2019, **364**.

143 A. Gesell, M. Rolf, J. Rg Ziegler, M. Luisa, D. Chá Vez, F.-C. Huang and T. M. Kutchan, 2009.

144 H. Hong, Y. I. Lee and D. Jin, *Microchem. J.*, 2010, **96**, 374–379.

145 Y. Wang, University College London, 2021.

146 S.-C. Li, N. K. Goto, K. A. Williams, C. M. Deber and G. D. Fasman, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 6676–6681.

147 X. Chen, X. Zheng, S. Ali, M. Guo, R. Zhong, Z. Chen, Y. Zhang, H. Qing and Y. Deng, *ACS Chem. Neurosci.*, 2018, **9**, 1388–1398.

148 R. Roddan, G. Gygli, A. Sula, D. Méndez-Sánchez, J. Pleiss, J. M. Ward, N. H. Keep and H. C. Hailes, *ACS Catal.*, 2019, **9**, 9640–9649.

149 R. Roddan, A. Sula, D. Méndez-Sánchez, F. Subrizi, B. R. Lichman, J. Broomfield, M. Richter, J. N. Andexer, J. M. Ward, N. H. Keep and H. C. Hailes, *Commun. Chem.*, 2020, **170**, 1–10.

150 M. E. Welsch, S. A. Snyder and B. R. Stockwell, *Curr. Opin. Chem. Biol.*, 2010, **14**, 347–361.

151 K. Ishiguro, S. Sakiyama, K. Takahashi and T. Arai, *Biochemistry*, 1966, **119**, 566–3775.

152 I. P. Singh and P. Shah, *Expert Opin. Ther. Pat.*, 2017, **27**, 17–36.

153 M. A. Weber, J. M. Neutel and D. H. G. Smith, *Am. Heart J.*, 1992, **123**, 1414–1420.

154 K. Koketsu, A. Minami, K. Watanabe, H. Oguri and H. Oikawa, 2012.

155 Y. Ji, J. Wang, M. Chen, L. Shi and Y. Zhou, *Chinese J. Chem.*, 2018, **36**, 139–142.

156 M. Hestericová, T. Heinisch, L. Alonso-Cotchico, J. D. Maréchal, P. Vidossich and T. R. Ward, *Angew. Chemie Int. Ed.*, 2018, **57**, 1863–1868.

157 H. Li, P. Tian, J.-H. Xu and G.-W. Zheng, *Org. Lett.*, 2017, **19**, 3151–3154.

158 M. Nishihachijo, Y. Hirai, S. Kawano, A. Nishiyama, H. Minami, T. Katayama, Y. Yasohara, F. Sato and H. Kumagai, *Biosci. Biotechnol. Biochem.*, 2014, **78**, 701–707.

159 D. A. Evans, M. D. Ennis and D. J. Mathre, *J. Am. Chem. Soc.* , 1982, **104**, 1737–1739.

160 D. A. Evans, J. Bartroli and T. L. Shih, *J. Am. Chem. Soc.*, 1981, **103**, 2127–2129.

161 K. Nozaki, N. Sakai, T. Nanno, T. Higashijima, S. Mano, T. Horiuchi and H. Takaya, *J. Am. Chem. Soc.* , 1997, **119**, 4413–4423.

162 P. Galletti, E. Emer, G. Gucciardo, A. Quintavalla, M. Pori and D. Giacomini, *Org. Biomol. Chem.*, 2010, **8**, 4117.

163 United States Patent Office, US 2663709, 1953, 1.

164 T. H. Lee, M. Son and S. Y. Kim, *Bio. Pharm. Bull.* , 2010, **6**, 958–962.

165 A. M. Klibanov, *Nature*, 2001, **409**, 241–246.

166 P. Galletti, E. Emer, G. Gucciardo, A. Quintavalla, M. Pori and D. Giacomini, *Org. Biomol. Chem.*, 2010, **8**, 4117–4123.

167 Z.-F. Sun, L.-N. Zhou, T. Zhang and Z.-T. Du, *Chinese Chem. Lett.* , 2017, **28**, 558–562.

168 T. Itagaki, A. Kawamata, M. Takeuchi, K. Hamada, Y. Iwabuchi, T. Eguchi, F. Kudo, T. Usui and N. Kanoh, *J. Antibiot. (Tokyo).*, 2016, **69**, 287–293.

169 C. Gregg and M. V. Perkins, *Org. Biomol. Chem.*, 2012, **10**, 6547.

170 D. Perrin, *Dissociation constants of organic bases in aqueous solution*, Butterworths, London, 1965.

171 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 486–501.

172 W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 125–132.

173 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658–674.

174 G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 355–367.

175 D. Liebschner, P. V. Afonine, N. W. Moriarty, B. K. Poon, O. V. Sobolev, T. C. Terwilliger and P. D. Adams, *Acta Crystallogr. Sect. D Struct. Biol.*, 2017, **73**, 148–157.

176 D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2016, **79**, 629–661.

177 S. Sasidharan, Y. Chen, D. Saravanan, K. M. Sundram and L. Y. Latha, *Afr J Tradit Complement Altern Med*, 2011, **8**, 1–10.

178 P. S. Baran, *J. Am. Chem. Soc.*, 2018, **140**, 4751–4755.

179 M. D. Truppo, *ACS Med. Chem. Lett.*, 2017, **8**, 476–480.

180 M. E. Pyne, K. Kevvai, P. S. Grewal, L. Narcross, B. Choi, L. Bourgeois, J. E. Dueber and V. J. J. Martin, *Nat. Commun.*, 2020, **11**, 1–10.

181 M. L. Sethi, *J. Pharm. Sci.*, 1983, **72**, 538–541.

182 H. Q. Wang, J. Hu, H. Y. Yan, S. Wu and Y. H. Li, *J. Asian Nat. Prod. Res.*, 2017, **19**, 1124–1133.

183 L. Wu, W. Zhang, X. Qiu, C. Wang, Y. Liu, Z. Wang, Y. Yu, R. D. Ye and Y. Zhang, *Molecules*, 2018, **23**, 1–10.

184 H. L. Li, T. Han, R. H. Liu, C. Zhang, H. S. Chen and W. D. Zhang, *Chem. Biodivers.*, 2008, **5**, 777–783.

185 Y. S. Kwon and M. Son, *Biomol. Ther.*, 2013, **21**, 181–189.

186 B. Ribár, D. Lazar, P. Radivojević, P. Engel, O. Gasic and I. Kanyo, *Structure of Corydaline*, Oxford Univ. Press. UGOZZOLI, F, 1976, vol. 32.

187 D. S. Bhakuni, *Curr. Sci.*, 1983, **52**, 897–903.

188 G. Blashke, *Arch. Pharm. (Weinheim).*, 1968, **432**, 439–443.

189 H. L. Holland, M. Castillo, D. B. Maclean and I. D. Spenser, *Can. J. Chem.*, 1974, **52**, 2818–2831.

190 H. Lechner, P. Soriano, R. Poschner, H. C. Hailes, J. M. Ward and W. Kroutil, *Biotechnol. J.*, 2018, **13**.

191 M. Rueffer, W. Bauer, M. H. Zenk and M. Ruefeer, *Can. J. Chem.*, 1994, **72**, 170–175.

192 E. Späth and E. Kruta, *Berichte der Dtsch. Chem. Gesellschaft*, 1929, **62**, 1024–1029.

193 M. Cushman and F. W. Dekow, *Tetrahedron*, 1977, **34**, 1435–1439.

194 C. Saá, E. Guitián, L. Castedo, R. Suau and J. M. Saá, *J. Org. Chem.*, 1986, **51**, 2781–2784.

195 T. F. Knöpfel, P. Aschwanden, T. Ichikawa, T. Watanabe and E. M. Carreira, *Angew. Chemie - Int. Ed.*, 2004, **43**, 5971–5973.

196 S. Zhou and R. Tong, *Org. Lett*, 2017, **19**, 1594–1597.

197 F. Subrizi, Y. Wang, B. Thair, D. Méndez-Sánchez, R. Roddan, M. Cárdenas-Fernández, J. Siegrist, M. Richter, J. N. Andexer, J. M. Ward and H. C. Hailes, *Manuscr. Submitt.*, 2021.

198 B. Horst, M. J. Wanner, S. I. Jørgensen, H. Hiemstra and J. H. Van Maarseveen, *J. Org. Chem.*, 2018, **83**, 15110–15117.

199 D.-J. Chang, H. An, K.-S. Kim, H. H. Kim, J. Jung, J. M. Lee, N.-J. Kim, Y. T. Han, H. Yun, S. Lee, G. Lee, S. Lee, J. S. Lee, J.-H. Cha, J.-H. Park, J. W. Park, S.-C. Lee, S. G. Kim, J. H. Kim, H.-Y. Lee, K.-W. Kim and Y.-G. Suh, *J. Med. Chem.*, 2012, **55**, 10863–10884.

200 D. Webb and T. F. Jamison, *Org. Lett.*, 2012, **14**, 568–571.

201 A. W. Struck, M. L. Thompson, L. S. Wong and J. Micklefield, *ChemBioChem*, 2012, **13**, 2642–2655.

202 H. Schönherr and T. Cernak, *Angew. Chemie - Int. Ed.*, 2013, **52**, 12256–12267.

203 S. Mordhorst, J. Siegrist, M. Müller, M. Richter and J. N. Andexer, *Angew.*

Chemie Int. Ed., 2017, **56**, 4037–4041.

204 R. S. Reczkowski, J. C. Taylor and G. D. Markham, *Biochemistry*, 1998, **37**, 13499–13506.

205 Z. Fu, Y. Hu, G. Markham and F. Takusagawa, *J. Biomol. Struct. Dyn.*, 1996, **13**, 727–739.

206 J. Komoto, T. Yamada, Y. Takata, G. D. Markham and F. Takusagawa, *Biochemistry*, 2004, **43**, 1821–1831.

207 F. della Ragione, M. Porcelli, M. Carteni-Farina, V. Zappia and A. E. Pegg, *Biochem. J.*, 1985, **232**, 335–341.

208 S. Wang, J. Lim, K. Thomas, F. Yan, R. H. Angeletti and V. L. Schramm, *J. Am. Chem. Soc.*, 2012, **134**, 1468–1470.

209 J. Siegrist, S. Aschwanden, S. Mordhorst, L. Thöny-Meyer, M. Richter and J. N. Andexer, *ChemBioChem*, 2015, **16**, 2576–2579.

210 C. R. Creveling, N. Morris, H. Shimizu, H. H. Ong and J. Daly, *Mol. Pharmacol.*, 1972, **8**, 398–409.

211 J. Siegrist, J. Netzer, S. Mordhorst, L. Karst, S. Gerhardt, O. Einsle, M. Richter and J. N. Andexer, *FEBS Lett.*, 2017, **591**, 312–321.

212 J. T. Nelson, J. Lee, J. W. Sims and E. W. Schmidt, *Appl. Environ. Microbiol.*, 2007, **73**, 3575–3580.

213 B. J. C. Law, M. R. Bennett, M. L. Thompson, C. Levy, S. A. Shepherd, D. Leys and J. Micklefield, *Angew. Chemie - Int. Ed.*, 2016, **55**, 2683–2687.

214 K. D. McMurtrey, L. R. Meyerson, J. L. Cashaw and V. E. Davis, *J. Org. Chem.*, 1984, **49**, 947–948.

215 B. Daniel, B. Konrad, M. Toplak, M. Lahham, J. Messenlehner, A. Winkler and P. Macheroux, *Arch. Biochem. Biophys.*, 2017, 632, 88–103.

216 R. Suau, M. V. Silva and M. Valpuesta, *Tetrahedron*, 1991, **47**, 5841–5846.

217 US 2007/0249650 A1, 2007.

218 L. Grycová, J. Dostál and R. Marek, *Phytochemistry*, 2007, **68**, 150–175.

219 S.-T. Lu, T.-L. Su, T. Kametani, A. Ujiie, M. Ihara and K. Fukumoto, *Heterocycles*, 1975, **3**, 459–465.

220 M. U. S. J. C. Sot and P. Trans, *Phytochemistry*, 1982, **21**, 809–810.

221 S. Wolfe, H. B. Schlegel and M.-H. Whangbo, *Can. J. Chem.*, 1974, **52**, 3787–2792.

222 P. D. Baird, J. Blagg, S. G. Davies and K. H. Sutton, *Tetrahedron*, 1988, **44**, 171–186.

223 W. Wiegrebe, U. Krüger, H. Reinhart and F. L, *Arch. Pharm. (Weinheim).*, 1968, **345**, 50–58.

224 K. Y. Zee-Cheng, K. D. Pauli and C. C. Cheng, *J. Med. Chem.*, 1974, **17**, 347–351.

225 K. B. Choi, T. Morishige and F. Sato, *Phytochemistry*, 2001, **56**, 649–655.

226 M. R. Bennett, M. L. Thompson, S. A. Shepherd, M. S. Dunstan, A. J. Herbert, D. R. M. Smith, V. A. Cronin, B. R. K. Menon, C. Levy and J. Micklefield, *Angew. Chemie - Int. Ed.*, 2018, **57**, 10600–10604.

227 M. Bala, S. Kumar, K. Pratap, P. K. Verma, Y. Padwad and B. Singh, *Nat. Prod. Res.*, 2019, **33**, 622–627.

228 J. B. I. De La Peña, H. L. Lee, S. Y. Yoon, G. H. Kim, Y. S. Lee and J. H. Cheong, *J. Nat. Med.*, 2013, **67**, 814–821.

229 J. L. Collins, A. Fujii, S. Roshandel, C. A. To and M. P. Schramm, *Bioorganic Med. Chem. Lett.*, 2017, **27**, 2953–2956.

230 J. D. La Chapa, M. Valdez, F. Ruiz, K. Gonzales, W. Mitchell, S. F. McHardy, M. Hart, S. R. Polusani and C. B. Gonzales, *Bioorganic Med. Chem.*, 2019, **27**, 208–215.

231 P. Cheng, N. Huang, Z. Y. Jiang, Q. Zhang, Y. T. Zheng, J. J. Chen, X. M. Zhang and Y. B. Ma, *Bioorganic Med. Chem. Lett.*, 2008, **18**, 2475–2478.

232 US 3597431, 1971.

233 M. Ružič, A. Pečavar, D. Prudič, D. Kralj, C. Scriban and A. Zanotti-Gerosa, *Org.*

Process Res. Dev., 2012, **16**, 1293–1300.

234 H. Nie, Y. Zhu, X. Hu, Z. Wei, L. Yao, G. Zhou, P. Wang, R. Jiang and S. Zhang, *Org. Lett.*, 2019, **21**, 8641–8645.

235 M. W. Chen, Y. Ji, J. Wang, Q. A. Chen, L. Shi and Y. G. Zhou, *Org. Lett.*, 2017, **19**, 4988–4991.

236 A. Pictet and A. Gams, *Ber.*, 1910, **42**, 2384–2391.

237 J. R. Ludwig and C. S. Schindler, *Chem*, 2017, **2**, 313–316.

238 2011/077405 A1, 2011.

239 J. Zhu, H. Tan, L. Yang, Z. Dai, L. Zhu, H. Ma, Z. Deng, Z. Tian and X. Qu, *ACS Catal.*, 2017, **7**, 7003–7007.

240 P. N. Craig, F. P. Nabenhauer, P. M. WIlliams, E. Macko and J. Toner, *J. Am. Chem. Soc.*, 1952, **74**, 1316–1317.

241 C. Xu, K. Xu, H. Gu, R. Zheng, H. Liu, X. Zhang, Z. Guo and B. Xu, *JACS Commun.*, 2004, **126**, 9938–9939.

242 World Intellectual Property Organization, WO 2008/069632, 2008, 1–33.

243 A. Y. Robin, C. Giustini, M. Graindorge, M. Matringe and R. Dumas, *Plant J.*, 2016, **87**, 641–653.

244 WO 2019/137985 A1, 2019, 1–48.

245 J. P. O'Bryan, *Pharmacol Res.*, 2019, **139**, 503–511.

246 M. Maniscalco, D. Singh-Franco, W. R. Wolowich and R. Torres-Colón, *Clin. Ther.*, 2006, **28**, 1247–1272.

247 WO 2008/011462 A2, 2008.

248 T. Q. Hu, G. R. Cairns and B. R. James, *Holzforschung*, 2000, **54**, 127–132.

249 R. A. Miller, A. M. Gilbert, S. Xue and W. D. Wulff, *Synthesis (Stuttg).*, 1998, 80–84.

250 L. Evanno, J. Ormala and P. M. Pihko, *Chem. Eur. J.*, 2009, **15**, 12963–12967.

251 I. M. Khalil, D. Barker and B. R. Copp, *J. Org. Chem.*, 2016, **81**, 282–289.

252 W. A. Hendrickson, *Q. Rev. Biophys.*, 2014, **47**, 49–93.

253 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658–674.

254 A. Urzhumtsev, P. V. Afonine and P. D. Adams, *Crystallogr. Rev.*, 2013, **19**, 230–270.

255 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.

256 R. Roddan, G. Gygli, A. Sula, D. Méndez-Sánchez, J. Pleiss, J. M. Ward, N. H. Keep and H. C. Hailes, *ACS Catal.*, 2019.

257 H. Schönherr and T. Cernak, *Angew. Chemie Int. Ed.*, 2013, **52**, 12256–12267.

258 K. Iwasa, M. Moriyasu, Y. Tachibana, H.-S. Kim, Y. Wataya, W. Wiegrebe, K. F. Bastow, L. M. Cosentino, M. Kozuka and K.-H. Lee, *Bioorganic Med. Chem.*, 2001, **9**, 2871–2884.

259 A. Mishra, S. Singh and S. Shukla, *J. Exp. Neurosci.*, 2018, **12**, 1–8.

260 L. C. S. L. Morais, J. M. Barbosa-Filho and R. N. Almeida, *J. Ethnopharmacol.*, 1998, **62**, 57–61.

261 N. Kondo, A. Takahashi, K. Ono and T. Ohnishi, *J. Nucleic Acids*, 2010, **2010**, 1–7.

262 M. R. Bennett, S. A. Shepherd, V. A. Cronin and J. Micklefield, *Curr. Opin. Chem. Biol.*, 2017, **37**, 97–106.

263 J. F. Carpenter, *J. Org. Chem*, 1993, **58**, 1607–1609.

264 I. Giuffrè, L. Taverniti and S. Di Staso, *Eur. J. Pphthalmology*, 2004, **14**, 508–13.

265 M. F. Rousseau, J. Raigoso, C. van Eyll, H. Van Mechelen, N. R. Musso, G. Lotti and H. Pouleur, *J. Cardiovasc. Pharmacol.*, 1992, **19**, 155–62.

266 F. Horak, P. Zieglmayer, R. Zieglmayer, P. Lemell, R. Yao, H. Staudinger and M. Danzig, *Ann. Allergy, Asthma Immunol.*, 2009, **102**, 116–120.

267 A. Bonamore, I. Rovardi, F. Gasparrini, P. Baiocco, M. Barba, C. Molinaro, B. Botta, A. Boffi and A. Macone, *Green Chem.*, 2010, **12**, 1623.

268 E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch, *Proteomics Protoc. Handb.*, 2005, 571–608.

269 T. Pesnot, M. C. Gershater, J. M. Ward and H. C. Hailes, *Adv. Synth. Catal.*, 2012, **354**, 2997–3008.

270 A. V. Malkov, K. Vrankova, M. Cerny and P. Kocovsky, *J. Org. Chem.* , 2009, **74**, 8425–8427.

271 W. Qian, W. Lu, H. Sun, Z. Li, L. Zhu, R. Zhao, L. Zhang, S. Zhou, Y. Zhou, H. Jiang, X. Zhen and H. Liu, *Bioorg. Med. Chem.*, 2012, **20**, 4862–4871.

272 S. McNicholas and E. Potterton, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 386–394.

273 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminform.*, 2012, **4**.

274 A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.

275 K. Katoh, K. Misawa, K. I. Kuma and T. Miyata, *Nucleic Acids Res.*, 2002, **30**, 3059–3066.

276 S. B. Needleman and C. D. Wunsch, *J. Mol. Biol.*, 1970, **48**, 443–453.

277 C. Notredame, D. G. Higgins and J. Heringa, *J. Mol. Biol.*, 2000, **302**, 205–217.